10019071 .051502

JC10 Rec'd PCT/PTO 26 DEC 2001

FORM PTO-1390 (Modified) (REV 5-93) U S DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: 065691-0266

U S APPLICATION NO (If known, see 37 CFR 1.5): Unassigned 10/019071

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/FR00/01747	22 June 2000	22 June 1999

TITLE OF INVENTION
1CBP90 Polypeptide and its Fragments and Polynucleotides Coding for said Polypeptides and Applications for Diagnosing and Treating Cancer

APPLICANT(S) FOR DO/EO/US
Christian Bronner, Raphael Hopfner, Marc Mousli, Jean-marc Jeltsch, Yves Lutz, and Pierre Oudet

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   - ☒ is transmitted herewith (required only if not transmitted by the International Bureau).
   - ☐ has been transmitted by the International Bureau.
   - ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)
6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   - ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   - ☐ have been transmitted by the International Bureau.
   - ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   - ☒ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).
11. ☐ Applicant claims small entity status under 37 CFR 1.27 .

Items 12. to 17. below concern other document(s) or information included:

12. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
13. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
14. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
15. ☐ A substitute specification.
16. ☐ A change of power of attorney and/or address letter.
17. ☒ Other items or information: Application Data Sheet

531 Rec'd PCT/PTO 26 DEC 2001

U.S. APPLICATION NO (if known, see 37 C.F.R 1.50) Unassigned 10/019071	INTERNATIONAL APPLICATION NO PCT/FR00/01747	ATTORNEY'S DOCKET NUMBER 065691-0266

18. ☒ The following fees are submitted:	CALCULATIONS	PTO USE ONLY
Basic National Fee (37 CFR 1.492(a)(1)-(5): Search Report has been prepared by the EPO or JPO $890.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482) $710.00		
No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2) $740.00		
Neither international preliminary examination fee (37 CFR 1.482) nor International search fee (37 CFR 1.445(a)(2)) paid to USPTO $1,040.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) $100.00		
ENTER APPROPRIATE BASIC FEE AMOUNT =	$890.00	
Surcharge of $130.00 for furnishing the oath or declaration later than 20 Months from the earliest claimed priority date (37 CFR 1.492(e))		

Claims	Number Filed		Included in Basic Fee		Extra Claims		Rate		
Total Claims	20	-	20	=	0	×	$18.00	$0.00	
Independent Claims	3	-	3	=	0	×	$84.00	$0.00	
Multiple dependent claim(s) (if applicable)							$280.00		
TOTAL OF ABOVE CALCULATIONS =								$890.00	
Reduction by ½ for filing by small entity, if applicable.								$0.00	
SUBTOTAL =								$890.00	
Processing fee of $130.00 for furnishing English translation later the 20 months from the earliest claimed priority date (37 CFR 1.492(f). +									
TOTAL NATIONAL FEE =								$890.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). $40.00 per property +									
TOTAL FEES ENCLOSED =								$890.00	

				Amount to be: refunded	$
				charged	$

a. ☒ A check in the amount of $890.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. 19-0741 in the amount of $0.00 to the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 19-0741. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Foley & Lardner
Customer Number: 22428

22428

PATENT TRADEMARK OFFICE

Phillip J. Articola

SIGNATURE

Phillip J. Articola

NAME for STEPHEN B. MAEBIUS

REGISTRATION NUMBER 35,264 Reg. No. 38,819

Rec'd PCT/PTO 15 MAY 2002

#4

10/019071

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Attorney Docket No. 065691-0266

Applicant:	BRONNER *et al.*
Title:	ICBP90 POLYPEPTIDE AND ITS FRAGMENTS AND POLYNUCLEOTIDES CODING FOR SAID POLYPEPTIDES AND APPLICATIONS FOR DIAGNOSING AND TREATING CANCER
Appl. No.:	10/019,071
Filing Date:	December 26, 2001
Examiner:	Unknown
Art Unit:	Unknown

PRELIMINARY AMENDMENT

Commissioner for Patents
Box PATENT APPLICATION
Washington, D.C. 20231

Sir:

Prior to examination of the present Continuing Application, Applicant respectfully requests that the application be amended as follows:

In the Specification:

1. Please delete the second full paragraph starting at page 40, lines 15-30, and replace it with the following paragraph:

Briefly, the following oligonucleotides have been synthesized:

5'-AATTC<u>G**ATTGG**TTCT</u>G**ATTGG**TTCTG**ATTGG**TTCTT-3' (SEQ ID NO:13) and 5'-CTAGAAGAA**CCAAT**CAGAA**CCAAT**CAGAA**CCAAT**CG-3' (SEQ ID NO:14). These nucleotides were then hybridised. According to the documentation of the manufacturer (Clontech, Palo Alto, CA), the reporter construct targeted possesses three copies in tandem of the ICB2 sequence (ICB2X3). As mentioned above, one copy of ICB2 is underscored and the CCAAT sequences are in bold. To determine the specificity of protein

binding to the ICB box, the following oligonucleotides, containing three copies in tandem of the GC1 box (GC1X3), also present in the promoter, have been synthesized and hybridised:

5'-AATTC**GGGGCGGGG**CC**GGGGCGGG**CCC**GGGGCGGG**GCT-3' (SEQ ID NO:15)

5'-CTAGAG**CCCCGCCCC**GG**CCCCGCCCC**GG**CCCCGCCCC**GG-3' (SEQ ID NO:16)

2. Please delete the second full paragraph starting at page 42, lines 19-31, and replace it with the following paragraph:

To test the ability of the 59 kDa protein to bind specifically to the ICB1 and/or ICB2 boxes, three tandem copies of ICB2 (ICB2X3, sequences described above) were labelled at the terminal end with 32 P phosphore using the T4 polynucleotide kinase (New England Biolabs) and [λ^{32}P]ATP (160 mCi/mmol, ICN Irvine, CA, USA). To examine the specificity of the binding, oligonucleotides containing only one copy of the CCAAT box were synthesized:

ICB1: 5'-AGTCAGGG**ATTGG**CTGGTCTG-' (SEQ ID NO:17);

5'-CAGACCAG**CCAAT**CCCTGACT-3' (SEQ ID NO:18)

ICB2: 5'-AAGCTACG**ATTGG**TTCTTCTG-3' (SEQ ID NO:19);

5'-CAGAAGAA**CCAAT**CGTAGCTT-3' (SEQ ID NO:20).

3. At page 42, lines 32-33 to page 43, lines 1-19, please delete the entire paragraph and replace it with the following paragraph:

The ICBP-59 protein (1 μg) was incubated with 1 ng of oligonucleotide and labelled at its terminal end by phosphorous ^{32}P in 12% glycerol, 12 mM HEPES-NaOH (pH 7.9), 60 mM KC1, 4 mM Tris-HCl (pH 7.9), 100 ng BSA, 0.6 mM DTT, and 100 ng poly(dI/dC) in 20 μl (Inouye *et al.*, 1994). After a 30-minutes incubation at room temperature, the reaction mix was loaded in 6% polyacrylamide gels. In competition experiments, the quantity indicated of non-labelled oligonucleotides were added to the reaction mix 10 minutes before the addition of proteins. To examine the binding properties of ICBP90 with regard to the ICB2 box, we used the same protocol except that labelled oligonucleotide contained only one copy of the CCAAT sequence as described below:

ICB2: 5'-ATAAAGGCAAGCTACG**ATTGG**TTCTTCTGGACGGAGAC-3' (SEQ ID NO:21).

5'-GTCTCCGTCCAGAAGAA**CCAAT**CGTAGCTTGCCTTTAT-3' (SEQ ID NO:22).

Binding specificity was studied using a non-labelled nucleotide containing a GC box of the human topoisomerase IIa promoter:

5'-GAATTCGAGGGTAAA**GGGGCGGGG**TTGAGGCAGATGCCA-3' (SEQ ID NO:23).

5'-TGGCATCTGCCTCAA**CCCCGCCCC**TTACCCTCGAATTC-3' (SEQ ID NO:24).

Please insert the Sequence Listing filed concurrently herewith following page 77 of the original application and renumber pages 1-14 of the Sequence Listing as pages 78-91. Please also delete the original Sequence Listing filed with the application.

REMARKS

Applicants submit this Amendment to insert the required references to SEQ ID NOS of the Sequence Listing filed concurrently herewith, to delete the Sequence Listing filed previously, and to indicate the insertion point for the Sequence Listing filed concurrently herewith. Applicants respectfully request examination on the merits of this application. Receipt of the initial Office Action on the merits is awaited.

The above amendments do not constitute new matter. Accordingly, claims 1-46 are presented for the examination on the merits. Favorable consideration of the application, as amended, is earnestly requested.

The Commissioner is hereby authorized to charge any additional fees which may be required regarding this application under 37 C.F.R. §§ 1.16-1.17, or credit any overpayment, to Deposit Account No. 19-0741.

Respectfully submitted,

Date May 15, 2002

By Reg. No. 48,627

for Stephen B. Maebius
Attorney for Applicants
Registration No. 35,264

FOLEY & LARDNER
3000 K Street, N.W., Suite 500
Washington, D. C. 20007-5109
Telephone: (202) 672-5569
Facsimile: (202) 672-5399

MARKED UP VERSION SHOWING CHANGES MADE

In the Specification:

Marked up version of specification starting at page 1, paragraph 4, lines 29-35:

1. Please delete the second full paragraph starting at page 40, lines 15-30, and replace it with the following paragraph:

Briefly, the following oligonucleotides have been synthesized: 5'-AATTCG**ATTGG**TTCTG**ATTGG**TTCTG**ATTGG**TTCTT-3' (SEQ ID NO:13) and 5'-CTAGAAGAA**CCAAT**CAGAA**CCAAT**CAGAA**CCAAT**CG-3' (SEQ ID NO:14). These nucleotides were then hybridised. According to the documentation of the manufacturer (Clontech, Palo Alto, CA), the reporter construct targeted possesses three copies in tandem of the ICB2 sequence (ICB2X3). As mentioned above, one copy of ICB2 is underscored and the CCAAT sequences are in bold. To determine the specificity of protein binding to the ICB box, the following oligonucleotides, containing three copies in tandem of the GC1 box (GC1X3), also present in the promoter, have been synthesized and hybridised:

5'-AATTC**GGGGCGGG**GCC**GGGGCGGG**CCC**GGGGCGGG**GCT-3' (SEQ ID NO:15)

5'-CTAGAG**CCCCGCCCC**GG**CCCCGCCCC**GG**CCCCGCCCC**GG-3' (SEQ ID NO:16)

2. Please delete the second full paragraph starting at page 42, lines 19-31, and replace it with the following paragraph:

To test the ability of the 59 kDa protein to bind specifically to the ICB1 and/or ICB2 boxes, three tandem copies of ICB2 (ICB2X3, sequences described above) were labelled at the terminal end with 32 P phosphore using the T4 polynucleotide kinase (New England Biolabs) and [λ^{32}P]ATP (160 mCi/mmol, ICN Irvine, CA, USA). To examine the specificity of the binding, oligonucleotides containing only one copy of the CCAAT box were synthesized:

ICB1: 5'-AGTCAGGG**ATTGG**CTGGTCTG-' (SEQ ID NO:17);

5'-CAGACCAG**CCAAT**CCCTGACT-3' (SEQ ID NO:18)

ICB2: 5'-AAGCTACG**ATTGG**TTCTTCTG-3' (SEQ ID NO:19);

5'-CAGAAGAA**CCAAT**CGTAGCTT-3' (SEQ ID NO:20).

3. At page 42, lines 32-33 to page 43, lines 1-19, please delete the entire paragraph and replace it with the following paragraph:

The ICBP-59 protein (1 μg) was incubated with 1 ng of oligonucleotide and labelled at its terminal end by phosphorous ^{32}P in 12% glycerol, 12 mM HEPES-NaOH (pH 7.9), 60 mM KC1, 4 mM Tris-HCl (pH 7.9), 100 ng BSA, 0.6 mM DTT, and 100 ng poly(dI/dC) in 20 μl (Inouye *et al.*, 1994). After a 30-minutes incubation at room temperature, the reaction mix was loaded in 6% polyacrylamide gels. In competition experiments, the quantity indicated of non-labelled oligonucleotides were added to the reaction mix 10 minutes before the addition of proteins. To examine the binding properties of ICBP90 with regard to the ICB2 box, we used the same protocol except that labelled oligonucleotide contained only one copy of the CCAAT sequence as described below:

ICB2: 5'-ATAAAGGCAAGCTACG**ATTGG**TTCTTCTGGACGGAGAC-3' (SEQ ID NO:21).

5'-GTCTCCGTCCAGAAGAA**CCAAT**CGTAGCTTGCCTTTTAT-3' (SEQ ID NO:22).

Binding specificity was studied using a non-labelled nucleotide containing a GC box of the human topoisomerase IIa promoter:

5'-GAATTCGAGGGTAAA**GGGGCGGGG**TTGAGGCAGATGCCA-3' (SEQ ID NO:23).

5'-TGGCATCTGCCTCAA**CCCCGCCCC**TTACCCTCGAATTC-3' (SEQ ID NO:24).

002.830256.1

10019071 .051502
Rec'd PCT/PTO 15 MAY 2002
10/019071

ICBP90 POLYPEPTIDE AND ITS FRAGMENTS AND POLYNUCLEOTIDES CODING FOR SAID POLYPEPTIDES AND APLLICATIONS TO THE DIAGNOSIS AND TREATMENT OF CANCER

The present invention relates to a new ICBP90 polypeptide and its fragments, to the cloning of cDNA and polynucleotides coding for said polypeptides, to cloning and/or expression vectors including said polynucleotides, cells transformed by said vectors and specific antibodies directed against said polypeptides. The invention also relates to methods and kits for diagnosing cancers, to a method and kit for screening ligands of the polypeptides of the invention and of compounds which may be used as a drug for prevention and/or treatment of cancers.

DNA topoisomerases are highly preserved nuclear proteins during evolution, the main role of which is for controlling DNA conformation and topology in the nucleus, which are constantly altered by the various biological processes involving DNA such as for example, transcription and replication. Topoisomerases exert their action by cutting DNA and linking these lesions after having achieved the adequate conformational change.

In mammals and humans in particular, today, there are at least five different genes coding for a topoisomerase and at least two additional pseudogenes (for a review, see Nitiss 1998). Thus, topoisomerase I, coded by the TOP1 gene removes the superturns present in DNA while only cutting a single strand. Both topoisomerases of type II existing in humans called TopIIα and TopIIβ, alter DNA topology by introducing transient double strand cleavages (for a review, see

Wang 1996). Finally, there are two topoisomerases of type III coded by two localized genes in 17p11.2-12 and 22q11-12 and they only act against negative superturns of DNA.

In tumoral cells, topoisomerases of type II play a very important role; in these growing and rapidly dividing cells, there is a large need for maintaining DNA molecules in a proper conformation as high transcription and replication rates are required. Thus, the rates for topoisomerase II are generally higher in human tumoral cells than in normal tissues of the same origin. However, the high expression rate of topoisomerase IIα in tumoral cells may vary among two tumors of different natures affecting a same tissue. For example, the nucleus of cells from small cell carcinomas of the lung has a higher rate of topoisomerase IIα than the nucleus of cells from lung carcinomas with normal sized cells (Guinee *et al.*, 1996). In the same way, the rate of topoisomerase IIα in A59 cells is three times higher than in PC3 cells, both of these cell lines stemming from the adenocarcinoma of lung epithelium (Yamasaki *et al.*, 1996).

These observations suggest that topoisomerase IIα may be considered as a marker of cell proliferation for certain types of cancer. As the cancerous process is characterized by abnormal cell proliferation partly due to the loss of contact inhibition, topoisomerase IIα therefore appears as a preferential target for chemiotherapeutical drugs for treating cancer (Pommier *et al.*, 1994), and the present anticancer treatments largely resort to inhibitors of topoisomerases.

Most of these inhibitors exert their cytotoxic effects by stabilizing the DNA cleavage complex. Drugs like anthracyclines [doxorubicin (adriamycin) or epipodophyllotoxins (such as etoposide (VP-16) or teniposide (VM26))], acridines (such as mAMSA) and anthracendiones (for example, mitoxantrone) are examples of drugs which inhibit topoisomerases II which stabilize the cleavage complex. More recently, a new class of inhibitors of topoisomerases II has been developed; these inhibitors act at the level of catalytic activity and no longer by stabilizing the cleavage complex. The drug, fostriecin is an example of one of them (Boritzki *et al.*, 1988). Today these different drugs are used in palliative and curative anticancer treatments.

Nevertheless, one of the major problems encountered in the present anticancer treatments using inhibitors of topoisomerases is the emergence of a resistance to drugs (Kubo *et al.*, 1995). These resistances are either the occurrence of an overexpression of pumps providing efflux of drugs outside the cells before they reach their target (for example; P-glycoprotein, a protein associated with multi-drug resistance (MRP)), or the occurrence of a change in the expression rate of topoisomerase IIα (Deffie *et al.*, 1989; Fry *et al.*, 1991), or either both occurrences (for a review, see Isaacs *et al.*, 1998).

One of the aspects of the present invention is therefore to understand the regulatory mechanisms of the expression of the gene of topoisomerase IIα, in order to develop an alternative to the phenomenon of resistance to drugs, observed for certain cancers and this with the aim of enhancing the curative and

preventive treatment of cancers.

There are two types of type II topoisomerase which differ in their expression profile; topoisomerase IIα (Top IIα) (170 kD), essentially located in the nucleoplasm at the centromer of the mitotic chromosomes, participates in the fundamental biological processes which are replication, condensation of chromosomes and transcription. It seems that topoisomerase IIβ (Top II (180 kD) is rather involved in the transcription of ribosomal RNA, given the nucleolar localization of this enzyme. Both human type II topoisomerases are localized on two different chromosomes (17q21-22 for topoisomerase IIα and 3p24 for topoisomerase IIβ) (Tsai-Plugfelder *et al*, 1988; Drake *et al.*, 1989; Chung *et al.*, 1989; Jenkins *et al.*, 1992; Austin *et al.*, 1993).

Unlike topoisomerase IIβ, the expression of which is characterized by a relative consistency, topoisomerase IIα has a variation of expression depending on the proliferation state of cells and on their position in the cell cycle. Expression of messenger RNA (RNAm) is higher in proliferating cells than in arrested cells in confluence. The expression of topoisomerase IIα increases during the S phase of the cell cycle, reaching a maximum at the end of phase G2/M (Goswami *et al.*, 1996), the level of messenger RNA being ten times higher at the end of phase S than during phase G1. Also, there seems to be a coupling between the synthesis and degradation of topoisomerase IIα and chromosomal condensation/decondensation(Heck *et al.*, 1988).

Present knowledge concerning control of the gene

of topoisomerase IIα, all in all, remains rather scanty. Recently, a promoter region of about 650 base pairs has been described by Hockhauser *et al.* (1992), it has all the characteristics of a domestic gene, an absence of TATA box and a moderate content of GC sites (notably the presence of a Sp1 box which may replace the TATA box) are two examples of this. The presence of 5 inverted CCAAT boxes or ICBs is another feature of this type of promoter.

Transcription factors interacting with the promoter of the gene of human topoisomerase IIα have been described; c-myb (Brandt *et al.*, 1997), p53 (Sandri *et al.*, 1996), ATF (Lim *et al.*, 1998), Sp1 and Sp3 (Kubo *et al.*, 1995) may be mentioned. Whatever the case, apart from NF-Y (also called CBF, ACF and CP1, references in Isaacs *et al.*, 1996), the transcription factors which act on the ICB sequences of the promoter for the gene of human topoisomerase IIα have not yet been identified and characterized; Herzog and Zwelling (1997) have however revealed two proteins with an apparent molecular weight of 90 kD and 140 kD which bind ICB1 to ICB4 and ICB5, respectively. Isaacs and his collaborators (1996) have suggested that NFY as well as another unidentified protein recognize an ICB box of the promoter region of the gene of topoisomerase IIα; they have also shown that ICB2 mutations completely suppressed the reduction in promoter activity normally observed in cells arrested in confluence (Isaac et al., 1996). They identified NFY as a component of a complex induced by the proliferation and which binds *in vitro* to the ICN2 sequence of the promoter of the gene of human topoisomerase IIα,

although NF-Y is always detectable in cells arrested in confluence (Isaacs et al., 1996). They suggested that ICB2 acts as a negative regulator of the promoter of the gene of topoisomerase IIα of cells arrested in confluence and that this repression may be suppressed in proliferative cells. The ICB2 box of the promoter of the gene of topoisomerase IIα therefore plays a primordial role in the arrest of the normal proliferative process when the cells reach confluence.

Transcription factors binding to the ICB sequence as well as the ICB sequence itself therefore form molecular targets for controlling the expression rate of topoisomerase IIα. By intervening on these factors, controlling the expression of the gene of topoisomerase IIα and cell proliferation consequently may be contemplated.

The object of the present invention is to detect new transcription factors binding to the ICB box involved in the control of cell proliferation.

A recent technique called a "simple hybrid" system has been used, which allows DNAc clones coding for the proteins binding to this specific DNA of certain sequences to be isolated. This system has a double advantage as it is able not only to reveal DNA-protein interaction *in vivo* in yeast, but also to give direct access to complementary DNAs (cDNA) coding for the candidate proteins having a transcription factor activity. This system is mainly based on the construct of a test yeast strain according to the principle developed by Wang and Reed (1993). This yeast strain enables DNAc banks to be screened by demonstrating DNA-protein interaction *in vivo* through activation of a

reporter gene integrated within the genome of the test yeast.

The object of the present invention is therefore an isolated polypeptide designated as ICBP90 (inverted CCAAT box binding protein) with the amino acid sequence SEQ ID No.2. This sequence comprises:

a) a "ubiquitin" domain comprising the sequence of amino acids 1-75 of sequence SEQ ID No.2;

b) a "zinc finger" domain of the C4HC3 type comprising the sequence of amino acids 310-366 of sequence SEQ ID No. 2 and a "zinc finger" domain of the C3HC4 type comprising the sequence of amino acids 724-763 of sequence ID No.2;

c) a presumed "zipper leucine" domain comprising the sequence of amino acids 58-80 of sequence SEQ ID No.2;

d) two potential nuclear localization domains comprising the sequences of amino acids 581-600 and 648-670 of sequence SEQ ID No.2;

e) a site for phosphorylation with a tyrosine kinase comprising the sequence of amino acids 452-458 of sequence SEQ ID No.2;

f) sites for phosphorylation with a dependent cAMP/cGMP protein kinase comprising the sequences of amino acids 246-249, 295-298 and 648-651 of sequence SEQ ID No.2;

g) sites for phosphorylation with a casein kinase II comprising the sequence of amino acids 23-36, 57-60, 91-94, 109-112, 165-168, 265-268, 354-357 and 669-672 of sequence SEQ ID No.2;

h) sites for phosphorylation with a protein kinase C comprising the sequence of amino acids 82-84, 104-106, 160-162, 173-175, 251-253, 301-303, 380-382,

393-395, 504-506, 529-531, 625-627 and 639-641 of sequence SEQ ID No.2.

The present invention also relates to an isolated polypeptide characterized in that, it comprises a polypeptide selected from:

a) a polypeptide of sequence SEQ ID No.2, SEQ ID No.4, SEQ No.6 or SEQ ID No.8;

b) a polypeptide, a polypeptide variant of sequences of amino acids defined under a);

c) a polypeptide homologous to the polypeptide defined under a) or b) and including at least 80% homology, preferably 90% with said polypeptide of a);

d) a fragment of at least 5 consecutive amino acids of a polypeptide defined under a), b) or c);

e) a biologically active fragment of a polypeptide defined under a), b) or c).

It should be understood that the invention relates to polypeptides obtained through purification from natural sources or else obtained through genetic recombination or even by chemical synthesis and they may then include non natural amino acids.

In the present specification, the term "polypeptide" will be used for also designating a protein or a peptide.

The term "polypeptide variant" shall be understood as designating all the mutated polypeptides which may exist in nature, in particular in the human being, and which notably correspond to truncations, substitutions, deletions and/or additions of amino acid residues. The homologous polypeptides according to the invention at least retain a domain selected from the DNA binding domain and/or the interaction domain with another protein.

It shall be understood that the term "homologous polypeptide" designates polypeptides having certain modifications, as compared with the natural polypeptide ECBP90, as in particular a deletion, addition or substitution of at least one amino acid, a truncation, an extension and/or a chimeric fusion. Among the homologous polypeptides, those for which the sequence of amino acids have at least 80% homology, preferably 90%, more preferably 95%, and most preferably 97% homology with the sequences of amino acids of the polypeptides according to the invention, are preferred. In the case of a substitution, one or several consecutive or non consecutive amino acids are replaced with "equivalent" amino acids. Here, the expression "equivalent" amino acid aims at designating any amino acid capable of being substituted for one of the amino acids of the basic structure without however changing the essential functional properties or characteristics, such as their biological activities, of the corresponding polypeptides such that induction *in vivo* of antibodies capable of recognizing the polypeptide for which the amino acid sequence is comprised within the amino acid sequence SEQ ID No.2, or in one of its fragments as defined above, and notably the sequence of amino acids SEQ ID No.4, SEQ ID No.6 and SEQ ID No.8. These equivalent amino acids may be determined either by relying on their structural homology with the amino acids which they replace, or on the results of cross biological activity tests which may take place for the different polypeptides. As an example, the possibilities of substitutions which may carried out without their resulting a deep change in the biological activities of the corresponding modified polypeptides

will be mentioned, for example replacements of leucine with valine or isoleucine, of aspartic acid with glutamic acid, of glutamine with asparagine, of arginine with lysine etc., the reverse substitutions may naturally be contemplated under the same conditions.

It shall be understood that the term "biologically active fragment" designates in particular a fragment of an amino acid sequence of a polypeptide according to the invention having at least one of the functional characteristics or properties of the polypeptides according to the invention, notably in that: (i) it is capable of being recognized by a specific antibody of a polypeptide according to the invention; (ii) it has at least one of the domains or regions as defined above; (iii) it is capable of binding to DNA and notably to the CCAATT and/or inverted CCAAT boxes; (iv) it is capable of modulating the expression rate of the gene of topoisomerase IIα, (v) it is capable of modulating cell proliferation.

It is understood that the term "polypeptide fragment" designates a polypeptide including a minimum of 5 amino acids, preferably 7 amino acids, more preferably 10, and most preferably 15 amino acids. Fragments of a polypeptide according to the invention, obtained by cleaving said polypeptide with a proteolytic enzyme, with a chemical reagent, or even by placing said polypeptide in a very acid environment, are also part of the invention.

The polypeptide according to the invention may also be associated with other polypeptides through protein-protein interactions. It is understood that the term "protein-protein interactions" designate

associations which directly bring into contact at least two proteins. Thus, the polypeptide of the invention may dimerize in order to form homodimers or heterodimers, or be associated as homomultimers or heteromultimers. The polypeptide according to the invention may also interact with another polypeptide in order to exert its action; hence, the polypeptide according to the invention may also have, in addition to its DNA binding domain, a domain acting on the transcription which exerts its action via protein-protein interactions with other protein components of the transcriptional machinery. It is understood that the term "protein component of the transcriptional machinery" designates all transcription factors required for performing and controlling the transcription reaction.

The polypeptide according to the invention is characterized in that it is capable of binding to a DNA sequence and in that it includes at least a DNA binding domain selected from the group consisting of a "zinc-finger" domain and a "leucine zipper" domain; the DNA sequence to which binds said polypeptide is a CCAAT box, preferably an inverted CCAAT box: ICB.

It is understood that the term "binding to a DNA sequence", designates a specific interaction between the polypeptide of the invention and a DNA sequence by means of a series of weak bonds formed between the amino acids of the protein and the bases. The polypeptide according to the invention, has at least a DNA binding domain which contains at least one of the known protein units capable of interacting with DNA, i.e. the zinc-finger structure with which is associated a zinc atom (zinc-finger) the helix-turn-helix

structure, the helix-loop-helix structure, and the leucine-zipper structure.

It is understood that the term "zinc-finger unit" designates a sequence of about twenty amino acids assuming a zinc-finger shape in space. There are two types of them: those which contain four cysteines (C4) and those which contain two cysteines and two histidines (C2H2). These amino acids define the nature of the zinc-finger and they are located at its base and a Zn^{++} ion is located in the middle of the square formed by these four amino acids. The polypeptide according to the invention potentially has two units of type C4.

It is understood that the term "leucine zipper type units" designates units belonging to dimeric transcription factors which are either homodimers or heterodimers. The monomer consists of a sequence with a basic character which interacts with DNA in a specific way and of a α helix hydrophobic domain which interacts with the homologous domain of the other chain. In this domain, leucine is found every 7 amino acids, i.e. at each turn of the helix. All these leucines are aligned and the interaction occurs at their level between both monomers. The polypeptide according to the invention potentially has a leucine zipper type unit.

The invention also relates to an isolated polynucleotide characterized in that it codes for a polypeptide of sequence SEQ ID No.1 as defined earlier. Preferably, the polynucleotide according to the invention has the SEQ ID No.1 sequence.

The invention also relates to the isolated polynucleotide characterized in that it comprises a polynucleotide selected from:

a) a polynucleotide with sequence SEQ ID No.1, SEQ ID No.3, SEQ ID No.5 or SEQ ID No.7 or for which the sequence is that of the RNA corresponding to sequence SEQ ID No1, SEQ ID No.3, SEQ No.5 or SEQ ID No.7;

b) a polynucleotide for which the sequence is complementary to the sequence of a polynucleotide defined under a),

c) a polynucleotide for which the sequence includes at least 80% homology with a polynucleotide defined under a) or b),

d) a polynucleotide which hybridizes under high stringency conditions with a polynucleotide sequence defined under a), b) or c),

e) a fragment of at least 15 consecutive nucleotides, preferably 21 consecutive nucleotides, and more preferably 30 consecutive nucleotides of a polynucleotide defined under a), b), c) or d), except for human EST AI084125, except for the sequence corresponding to sequence SEQ ID No.944 published on August 5th 1999 in Patent Application WO 99 38972 and except for sequences SEQ ID No.9, No.10 and No.11 corresponding to the human ESTs No. AI 0830773, No. AA 811055, No. AA 488 755, No. AA 129 794 and No. AA 354 253 present in the human EST data bases (human dbest), respectively.

In the present specification, it is understood that the terms, "polynucleotide, oligonucleotide, polynucleotide sequence, nucleotidic sequence, or nucleic acid", shall designate a DNA fragment, as well as a double strand DNA, a single strand DNA, as well as transcription products of said DNAs, and/or an RNA fragment, said isolated natural or synthetic fragments whether including non-natural nucleotides or not,

designating a specific chaining of nucleotides, whether modified or not, providing definition of a fragment or a region of a nucleic acid.

It is understood that the term "polynucleotide" with a complementary sequence, designates any DNA for which the nucleotides are complementary to those of SEQ ID No.1, SEQ ID No.3, SEQ ID No.5, SEQ ID No.7 or of a part of SEQ ID No.1, SEQ No.3, SEQ ID No.5, SEQ ID No.7 and for which the orientation is inverted.

In the sense of the present invention, it is understood that the term "homology percent" designates a percentage of identity between bases of two polynucleotides, this percentage being purely statistical and the differences between both polynucleotides are randomly distributed throughout their length. According to the invention, the polynucleotides with a homologous nucleic sequence have a homology rate of at least 80%, preferably 90%, more preferably 95%, most preferably 97%.

Hybridization under strong stringency conditions means that the temperature and ionic force conditions are selected in such a way that hybridization between two complementary DNA fragments may be maintained. As an illustration, strong stringency conditions of the hybridization step for the purpose of defining the polynucleotidic fragments described above, advantageously are the following:

DNA-DNA or DNA-RNA hybridization is achieved in two steps: (1) prehybridization at 42°C for 3 hours in phosphate buffer (20 mM pH 7.5) containing 5 x SSC (1 x SSC corresponds to a 0.15 M NaCl + 0.015 M sodium citrate solution), 50% formamide, 7% sodium dodecylsulfate (SDS), 10 x Denhard's, 5% dextran

sulfate and 1% salmon sperm DNA; (2) the actual hybridization for 20 hours at a temperature depending on the size of the probe (i.e. 42°C, for a probe with a size > 100 nucleotides), followed by two washings for 20 minutes at 20°C into 2 x SSC + 2% SDS, one washing for 20 minutes at 20°C into 0.1 x SSC + 0.1% SDS. The last washing is performed in 0.1 x SSC + 0.1% SDS for 30 minutes at 60°C for a probe with a size > 100 nucleotides. The strong stringency hybridization conditions described above, for a polynucleotide with a defined size, will be adapted by one skilled in the art for oligonucleotides with a larger or smaller size, according to the teaching of Sambrook *et al.*, 1989.

Advantageously, a nucleotidic fragment meeting the earlier definition will have at least 15 consecutive nucleotides, preferably at least 21 nucleotides, and even more preferably at least 30 consecutive nucleotides of the sequence from which it stems.

It is understood that the term EST ("expressed sequence tag") designates expressed sequences, characterized in a complementary DNA bank (DNAc) and used as a map marker for genomic DNA.

According to one embodiment of the invention, the polynucleotide according to the invention is characterized in that it is directly or indirectly labeled with a radioactive compound or a non-radioactive compound. Use of a polynucleotide according to the invention as a primer for amplifying or polymerizing nucleic sequences; the invention also relates to the use of a polynucleotide according to the invention as a probe for detecting nucleic sequences. According to the invention, the polynucleotide fragments may be used as a probe or as a primer in

methods for detecting, identifying, dosing and amplifying nucleic sequences, and they have a minimum size of 9 bases, preferably 18 bases, and more preferably 36 bases. Finally, the invention is related to the use of a polynucleotide according to the invention as a sense or anti-sense nucleic acid sequence for controlling the expression of the corresponding protein product.

The non-labeled sequences of polynucleotides according to the invention may directly be used as a probe, a primer or an oligonucleotide; however the used sequences are generally labeled for obtaining usable sequences for many applications. The labeling of primers, probes, oligonucleotides according to the invention is achieved through radioactive elements or through non-radioactive molecules; ^{32}P, ^{33}P, ^{35}S, ^{3}H, or ^{125}I may be mentioned among the used radioactive isotopes. The non-radioactive entities are selected from ligands such as biotin, avidin, streptavidin, dioxygenin, haptenes, dyes, luminescent agents, such as radioluminescent, chemiluminescent, bioluminescent, fluorescent, phosphorescent agents.

The polynucleotides according to the invention may thus be used as a primer and/or a probe in methods notably implementing the PCR (polymerase chain reaction) technique (Erlich, 1989; Innis *et al.*, 1990, and Rolfs *et al.*, 1991). This technique requires the selection of pairs of oligonucleotidic primers framing the fragment which should be amplified. Reference may for example, be made to the technique described in the US Patent No. 4,683,202. The amplified fragments may be identified, for example after agarose gel or polyacrylamide electrophoresis or after a

chromatographic technique like gel filtration or ion exchange chromatography. The specificity of the amplification may be controlled by molecular hybridization by using as a probe, nucleotidic sequences of polynucleotides of the invention, plasmids containing these sequences or their amplification products. Amplified nucleotidic fragments may be used as reagents in hybridization reactions in order to demonstrate the presence, in a biological sample, of a target nucleic acid with a sequence complementary to that of said amplified nucleotidic fragments.

The invention is also directed to nucleotidic fragments which may be obtained through amplification by means of primers according to the invention.

Other techniques for amplifying the target nucleic acid may advantageously be used as an alternative to PCR (PCR-like) by means of a pair of primers for nucleotidic sequences according to the invention. It is understood that the term "PCR-like" designates all methods implementing direct or indirect reproductions of nucleic acid sequences, or else those in which the labeling system has been amplified, of course these techniques are known, generally this deals with DNA amplification by a polymerase; when the original sample is an RNA, a reverse transcription should be performed beforehand. Presently, there are very many methods which provide such amplification, such as for example, the SDA (Strand Displacement Amplification) technique (Walker *et al.*, 1992), the TAS (Transcription-based Amplification System)technique described by Kwoh *et al.*, in 1989, the 3SR (Self-Sustained Sequence Replication) technique described by Guatelli *et al.*, in 1990, the NASBA (Nucleic Acid Sequence Based

Amplification) technique described by Kievitis *et al.*, in 1991, the TMA (Transcription Mediated Amplification) technique, the LCR (Ligase Chain Reaction) technique described by Landegren *et al.*, in 1988, and enhanced by Barany *et al.*, in 1991, which uses a thermostable ligase, the RCR (Repair Chain Reaction) technique described by Segev in 1992, the CPR (Cycling Probe Reaction) technique described by Duck *et al.*, in 1990, the Q-beta-replicase amplification technique described by Miele *et al.*, in 1983, and notably enhanced by Chu et al., in 1986 and Lizardi *et al.*, in 1988, and then by Burg *et al.*, as well as Stone *et al.*, in 1996.

If the target polynucleotide is an RNA, for example a RNAm, a reverse transcriptase type enzyme will advantageously be used before implementing an amplification reaction with the primers according to the invention or before implementing a detection method with probes of the invention, in order to obtain a DNAc from the RNA contained in the biological sample. The obtained DNAc will then be used as a target for the primers or the probes implemented in the detection or amplification method according to the invention.

The nucleotidic probes according to the invention, specifically hybridize with a DNA or RNA polynucleotide molecule according to the invention, more particularly with the sequence SEQ ID No.1 coding for the ECBP90 polypeptide, under strong stringency hybridization conditions such as those given as an example earlier.

The hybridization technique may be used in different ways (Matthews *et al.*, 1988). The most general method consists of immobilizing the nucleic acid extracted from cells of different tissues or from cells cultivated on a support (such as nitrocellulose,

nylon, polystyrene) and of incubating, under well defined conditions, the immobilized target nucleic acid with the probe. After hybridization, the probe excess is removed and the formed hybrid molecules are detected by the suitable method (measurement of radioactivity, fluorescence or enzyme activity related to the probe).

According to another embodiment of the nucleic probes, according to the invention, the latter may be used as a capture probe. In this case, a so-called "capture probe" is immobilized on a support and is used for capturing through specific hybridization, the target nucleic acid obtained from the biological sample to be tested and the target nucleic acid is then detected by a second probe, a so-called "detection probe", labeled with an easily detectable element.

In a preferred embodiment, the invention comprises the use of a sense or anti-sense oligonucleotide for controlling the expression of the corresponding protein product. Among the interesting nucleic acid fragments, anti-sense oligonucleotides i.e. those for which the structure provides an inhibition of the expression of the corresponding product, by hybridization with the target sequence, may be mentioned in particular. The sense oligonucleotides which, through interaction with the proteins involved in the control of the expression of the corresponding product which will induce either an inhibition, or an activation of this expression, should also be mentioned. The oligonucleotides according to the invention, have a minimum size of 9 bases, preferably 18 bases, and more preferably 36 bases.

The invention relates to a recombinant vector for cloning a polynucleotide according to the invention

and/or for expressing a polypeptide according to the invention characterized in that, it contains a polynucleotide according to the invention, as described earlier. The vector according to the invention, is characterized in that it includes components for the expression, possibly the secretion, of said sequences in a host cell. These vectors are useful for transforming host cells in order to clone or express nucleotidic sequences of the invention. Particular vectors are for examples the vectors of plasmidic or viral origin. Among these vectors, those of the pGEX series (Pharmacia) for expression in bacteria or pSG5 (Stratagene, La Jolla, CA USA) are preferred for expression in a eukaryotic system.

According to a particular embodiment, the vector according to the invention includes components for controlling expression of the polypeptides, these control components are preferably selected from (i) the promoter sequence of the ICBP90 gene according to the invention which corresponds to sequence SEQ ID No.12; (ii) a polynucleotide for which the sequence is complementary to the sequence SEQ ID No.12; (iii) a polynucleotide for which the sequence includes at least 80% identity with a polynucleotide as defined in (i) or (ii); (iv) a polynucleotide which hybridizes under strong stringency conditions with the polynucleotide sequence defined under (i), (ii), (iii). Computer tools available to one skilled in the art will easily allow him/her to identify the required and sufficient promoter control boxes for controlling the genic expression, notably the TATA, CCAAT, GC boxes, as well as enhancer or silencer control sequences which control in CIS the expression of genes according to the

invention.

The use of the above components defined and selected from the sequence SEQ ID No.12 for controlling the expression of heterologous polypeptides other than those of the invention and notably for controlling the expression of heterologous polypeptides in cell types in which the polypeptides according to the invention are expressed normally, is also within the scope of the invention.

The invention further comprises host cells, notably eukaryotic and prokaryotic cells, characterized in that they are transformed with vectors according to the invention. Preferably, the host cells are transformed under conditions allowing a recombinant polypeptide according to the invention to be expressed. The cell host may be selected from bacterial cells (Olins and Lee, 1993), but also from yeast cells (Buckholz, 1993), as well as animal cells, in particular mammal cell cultures (Edwards and Aruffo, 1993), but also insect cells wherein methods implementing baculoviruses for example may be used (Luckow, 1993). These cells may be obtained by introducing into the host cells a nucleotidic sequence inserted in a vector such as defined above, and then by growing said cells under conditions providing replication and/or expression of the transfected nucleotidic sequence.

The invention also relates to a method for preparing a polypeptide, characterized in that it implements a vector according to the invention. More specifically, the invention relates to a method for preparing a recombinant polypeptide characterized in that the transformed cells according to the invention

are grown under conditions providing expression of said recombinant polypeptide and in that said recombinant polypeptide is recovered.

The polypeptide according to the invention may be obtained according to a method of the invention, and according to production techniques for recombinant polypeptides, known to one skilled in the art. The present invention therefore relates to the recombinant polypeptide which may be obtained by the method shown above. In this case, the nucleic acid sequence used is placed under the control of signals providing its expression in a cell host. An efficient production system for a recombinant polypeptide requires the availability of a vector, for example of plasmidic or viral origin and of a compatible host cell. The vector should include a promoter, signals for initiating and terminating the translation, as well as suitable regions for controlling the transcription. It should be able to be maintained in the cell stably and may optionally have particular signals specifying the secretion of the translated polypeptide. These different control signals are selected depending on the used host cell. For this purpose, the nucleic acid sequences according to the invention may be inserted in autonomous replication vectors inside the selected host or integrative vectors of the selected host. Such vectors are prepared according to methods currently used by one skilled in the art and the resulting clones may be introduced into a suitable host by standard methods such as for example transfection with calcium phosphate precipitation, lipofection, electroporation, thermal shock.

The recombinant polypeptides obtained as indicated

above, may both exist in the glycosylated and non-glycosylated form and may have the natural tertiary structure or not.

The polypeptides obtained through chemical synthesis and which may include non-natural amino acids corresponding to said recombinant polypeptides, are also comprised in the invention. The peptides according to the invention may also be prepared by conventional techniques, in the field of peptide synthesis. This synthesis may be carried out in a homogenous solution or in the solid phase.

The methods used for purifying recombinant polypeptides are well known to one skilled in the art. The recombinant polypeptide may be purified from lysats and cell extracts, from the supernatant of the culture medium, by methods either used individually or in combination, such as fractionation, chromatography methods, immuno-affinity techniques by means of specific mono- or polyclonal antibodies, etc.

A preferred alternative consists of producing a recombinant polypeptide fusioned to a "carrier" protein (chimeric protein). The advantage of this system is that it provides stabilization and a reduction in the proteolysis of the recombinant product, an increase in the solubility during renaturation *in vitro* and/or a simplification of the purification when the fusion partner has an affinity for a specific ligand.

The invention also relates to a monoclonal or polyclonal antibody and to its fragments, characterized in that they specifically bind a polypeptide according to the invention. Chimeric antibodies, humanized antibodies and simple chain antibodies are also part of the invention. Antibody fragments according to the

invention are preferably Fab or F(ab')2 fragments.

The polypeptides according to the invention allow monoclonal or polyclonal antibodies to be prepared. Advantageously, monoclonal antibodies may be prepared from hybridomas according to the technique described by Kohler and Milstein in 1975. The inventors use this technique for obtaining a hybridoma producing a new highly specific monoclonal antibody of an epitope of protein ICBP90.

Polyclonal antibodies may be prepared, for example, by immunizing an animal, for example a mouse, with a polypeptide according to the invention associated with an adjuvant from the immune response, and then by purifying the specific antibodies contained in the serum of the immunized animals on an affinity column on which is fixed beforehand the polypeptide which has been used as an antigen. The polyclonal antibodies according to the invention may also be prepared by purification on an affinity column, on which a polypeptide according to the invention has been immobilized beforehand.

The invention also relates to a specific monoclonal antibody of the human ICBP90 protein and capable of inhibiting interaction between ICBP90 and the DNA sequence onto which protein ICBP90 specifically binds. According to another embodiment, the monoclonal antibody according to the invention and specific to the human ICBP90 protein is capable of inhibiting the interaction between ICBP90 and the proteins with which interacts ICBP90, said proteins preferably being ICBP90 itself, or proteins from the transcriptional complex. It is understood that the term "proteins from the transcriptional complex" designates all proteins

participating in the transcription reaction whether this happens in the initiation, elongation, or termination of the transcription.

The antibodies of the invention may also be labeled in the same way as described earlier for the nucleic probes of the invention, and preferably with an enzymatic, fluorescent or radioactive type labeling.

Moreover, in addition to their use for purifying polypeptides, the antibodies of the invention, in particular the monoclonal antibodies, may also be used for detecting these polypeptides in a biological sample.

They thus form a means for analyzing the expression of the polypeptide according to the invention, for example through immunofluorescence, labeling with gold, enzymatic immunoconjugates.

More generally, the antibodies of the invention may advantageously be implemented in any situation where the expression of a polypeptide according to the invention needs to be observed, and more particularly in immunocytochemistry, in immunohistochemistry, or in Western blotting experiments.

Thus, the invention relates to a method for detecting and/or dosing a polypeptide according to the invention, in a biological sample, characterized in that it comprises the following steps for bringing the biological sample into contact with antibodies according to the invention and then for detecting the formed antigen-antibody complex. This method may be used in immunocytochemistry for cell localization of the polypeptide according to the invention and in immunohistochemistry for assessing cell proliferation.

A kit for detecting and/or dosing a polypeptide

according to the invention in a biological sample, is also within the scope of the invention, characterized in that it comprises the following components: (i) a monoclonal or polyclonal antibody such as described earlier; (ii) if necessary, the reagents for forming the favorable medium for the immunological reaction; (iii) the reagents for detecting the antigen-antibody complexes produced by the immunological reaction. This kit is notably useful for conducting Western blotting experiments; with the latter, control of the expression of the polypeptide according to the invention may be investigated starting with tissues or cells. This kit is also useful for immunoprecipitation experiments in order to notably detect proteins which interact with the polypeptide according to the invention.

Any conventional procedure may be implemented for carrying out such a detection and/or dosage. As an example, a preferred method involves immunoenzymatic processes according to the immunofluorescence or radioimmunological (RIA) ELISA technique or equivalent.

The invention also comprises a method for detecting and/or dosing a nucleic acid according to the invention, in a biological sample, characterized in that it includes the following steps: (i) isolation of the DNA from the biological sample to be analyzed, or obtaining a DNAc from the RNA of a biological sample; (ii) specific amplification of the DNA coding for the polypeptide according to the invention by means of primers; (iii) analysis of the amplification products.

The invention further comprises a kit for detecting and/or dosing a nucleic acid according to the invention, in a biological sample, characterized in that it comprises the following components: (i) a pair

of nucleic primers according to the invention, (ii) the required reagents for carrying out a DNA amplification reaction and optionally (iii) a component for checking the sequence of the amplified fragment, more particularly a probe according to the invention.

The invention also comprises a method for detecting and/or dosing a nucleic acid according to the invention, in a biological sample, characterized in that it includes the following steps: (i) bringing a probe according to the invention into contact with a biological sample; (ii) detecting and/or dosing the hybrid formed between said probe and the DNA of the biological sample.

The invention also comprises a kit for detecting and/or dosing a nucleic acid according to the invention, in a biological sample, characterized in that it comprises the following components: (i) a probe according to the invention, (ii) the reagents required for implementing a hybridization reaction and if necessary, (iii) a pair of primers according to the invention, as well as the reagents required for an DNA amplification reaction.

The invention particularly relates to methods according to the invention and described above, for detecting and diagnosing cell proliferation, and more particularly cell proliferation of cancerous origin.

The invention also relates to a method for screening ligands able to affect the transcriptional activity of a gene, the promoter of which includes CCAAT and/or inverted CCAAT boxes capable of binding a polypeptide according to the invention, said method being characterized in that it includes the following steps for bringing into contact said polypeptide and

one or several potential ligands in the presence of reagents required for implementing a transcription or detection reaction and/or a reaction for measuring transcriptional activity. One of the objects of the invention is also to provide a kit or package for screening ligands able to affect the transcriptional activity of a gene, the promoter of which includes CCAAT and/or inverted CCAAT boxes capable of binding a polypeptide according to the invention characterized in that it comprises the following components: (i) a polypeptide according to the invention; (ii) a ligand; (iii) the reagents required for implementing a transcription reaction.

The ICBP90 polypeptide according to the invention has a nuclear receptor function. It is understood that the term "nuclear receptor" designates a polypeptide which has the essential properties of hormone nuclear receptors. This gene superfamily contains i.a. the retinoic acid nuclear receptors (RAR, RXR, ...), steroid hormone nuclear receptors (glucocorticoids, mineralocorticoids, progesterone, androgen, estrogen), and thyroid hormone nuclear receptors (T3 hormone). Accordingly, one of the objects of the present invention is also to provide a method for screening ligands able to affect the "nuclear receptor" function of the polypeptide according to the invention. Such a method includes the steps of:

a) bringing into contact the polypeptide of the invention and one or several potential ligands in the presence of required reagents;

b) detecting and/or measuring the transcriptional activity of a gene, the promoter of which includes nucleotidic sequences onto which the polypeptide of the

invention may be bound. Preferably, said nucleotidic sequences are CCAAT and/or inverted CCAAT boxes (ICB).

Techniques for detecting and/or measuring the transcriptional activity are known to one skilled in the art. The Northern blotting and RT-PCR technologies should notably be mentioned, which may be implemented with polynucleotides of the invention used as a probe or as a primer, respectively.

It is understood that the term "ligand" defines all compounds able to interact with the polypeptide according to the invention, in order to form a complex able to affect the transcriptional activity, i.e. to increase, reduce, modulate or cancel the transcription of a gene under the control of a promoter containing a DNA sequence to which binds the polypeptide of the invention.

Such a ligand is therefore able to have an agonist or antagonist activity. Among the ligands according to the invention, the biological molecules which interact with the polypeptide according to the invention as well as all the synthetic chemical compounds, should be mentioned. Among these ligands, the antibody according to the invention as well as an oligonucleotide having an identity of sequence with the CCAAT and/or inversed CCAAT nucleotidic sequence should also be mentioned; such a ligand is able to form an inhibitor of the polypeptide according to the invention.

The invention also relates to the ligand which may be obtained by the previous screening methods.

It is also understood that the term "ligand" defines any compound able to bind to the binding DNA sequence for the polypeptide according to the invention. Such a ligand forms a competitive inhibitor

of the polypeptide according to the invention for its binding to the DNA sequence.

Preferably, the biological sample according to the invention in which detection and dosage is performed, consists of a body fluid, for example human or animal serum, blood, saliva, lung mucus, or biopsies. The biological liquid resulting from a broncho-alveolar washing also obtained during analyses for diagnosing cancers of the deep airways is also included in the definition of a biological sample of the invention.

According to another aspect, the invention relates to a compound characterized in that it is selected from an antibody, a polypeptide, a ligand, a polynucleotide, an oligonucleotide, or a vector according to the invention as a drug, and notably as active ingredients of a drug: these compounds preferably will be in soluble form, associated with a pharmaceutically acceptable carrier. It is understood that the term "pharmaceutically acceptable carrier" designates any type of carrier usually used in preparing injectable compositions, i.e. a diluent, a suspension agent, such as an isotonic or buffered saline solution. Preferably, these compounds will be administered systemically, in particular intravenously, intramuscularly, intradermally, or orally. Their modes of administration, dosages and optimal dosage forms may be determined according to the criteria generally considered in establishing a suitable treatment for a patient as for example, the age or body weight of the patient, the seriousness of his/her general condition, tolerance to the treatment and ascertained secondary effects, etc.

According to another aspect, the invention relates

to a compound, characterized in that it is selected from a polypeptide, a polynucleotide, an anti-sense polynucleotide, an antibody, a vector, a cell, a ligand according to the invention as a drug and notably as active ingredients of a drug; these compounds preferably will be in soluble form, associated with a pharmaceutically acceptable carrier. It is understood that the term "pharmaceutically acceptable carrier" designates any type of carrier usually used in preparing injectable compositions, i.e. a diluent, a suspension agent, such as an isotonic or buffered saline solution. Preferably, these compounds are administered systemically, in particular intravenously, intramuscularly, intradermally or orally. Their modes of administration, dosages and optimal dosage forms may be determined according to criteria generally considered in establishing a suitable treatment for a patient such as for example the age or body weight of the patient, the seriousness of his/her general condition, tolerance to the treatment and the ascertained secondary effects, etc. When the agent is a polypeptide, an antagonist, a ligand, a polynucleotide, for example an anti-sense composition, a vector, it may be introduced into tissues or host cells by a number of ways, including viral infection, micro-injection, or fusion of vesicles. Jet injection for an intramuscular administration as described by Furth *et al.* (1992) may also be used. The polynucleotide may also be deposited on gold micro-particles, and be delivered intradermally by means of a particle bombardment apparatus, or a "gene pistol" as described in the literature (see for example Tang *et al.* (1992) where gold microprojectiles are coated with the polynucleotide of the invention,

preferably the anti-sense polynucleotide of the invention, then are bombarded into the skin cells.

The compound comprising this invention is used for the preparation of a pharmaceutic designed to modulate, raise, or diminish cellular proliferation.

The invention also has at its foundation a pharmaceutical composition that can act in the preventive and curative treatment of cancer and is characterised by a therapeutically effective quantity of an active compound and a pharmaceutically acceptable excipient. Using the preferred method of synthesis, this pharmaceutical composition contains antibodies that serve as targeting agents; those antibodies are conjugated to at least one agent selected from among antiproliferative, antineoplastic, or cytotoxic agents. These agents are either radioisotopes or non-isotopic substances. The conjugation of antibodies contained in the present invention with antiproliferative, antineoplastic, or cytotoxic agents can be utilized for arresting the development of cancers and for inducing regression and even elimination of tumoural masses. Preferably, the antibody or the antibody fragment conjugated to the agent is administered to the cancer patient and delivered to tumour sites by oral or parenteral route through a pharmaceutically acceptable transporting liquid, such as saline. Alternatively, a solution or suspension of antibody and antibody fragment conjugated to an agent can be perfused directly into the tissue of a malignant epithelial cancer, a method used by preference when the cancer has not metastasized.

For therapeutic use, the preferred radioisotopes, conjugated to monoclonal antibodies, are gamma

emitters, the most effective being iodine^{131}, yttrium^{90}, gold^{199}, palladium^{100}, copper^{67}, bismuth^{217}, and antimony^{211}. Alpha and beta emitting radioisotopes can also be employed for therapy. Non-isotopic substances conjugated to monoclonal antibodies and used for therapy are abundant and varied; for example: (i) antimetabolites, such as anti-folate agents like methotrexate, (ii) purine and pyrimidine analogues (mercaptopurine, fluorouracil, 5-azacytidine, (iii) antibiotics, (iv) lectins (ricin, abrin) and (iv) bacterial toxins (diphtheria toxin).

The antibodies of the invention can also be used as targeting agents to target cytotoxic cells, such as human T cells, monocytes or NK cells present or not at a metastasised tumour site. Antibodies can attach to cytotoxic cells via the Fc receptor situated at the surface of these cells or via an intermediary antibody that has a double specificity. Such bi-specific antibodies for the targeting of cancerous cells can be produced by fusing an immune cell producing the antibody of the present invention or a hybridoma of the present invention with a cell producing an antibody directed against the targeted cytotoxic cell. Bi-specific antibodies can equally be produced by chemically coupling two antibodies having the desired specificity. The antibodies of this invention also permit the targeting of carriers bearing antiproliferative, antineoplastic, or cytotoxic agents to the site of the tumor or metastatic tumor. By carriers we are referring to liposomes and viral particles. In certain cases, it's possible to predetermine the target elements to assure a specific expression in certain tissues or cells and limit the

expression zones of the polypeptides of this invention.

The invention also concern a product comprising at least a compound of the invention, and at least an anticancerous agent as a combination product for a simultaneous, separated or delayed use over the time.

In summary, the invention concerns a composition for the detection, localisation, and imaging of cancers, using an antibody that is tagged directly or indirectly by a marker whose signal is generated by radioactive or non-isotopic substances as defined above. The invention also has as objective the localisation and imaging of cancers, including (i) the stages of dispersion after parenteral injection into a human of a composition based on the invention; (ii) the accumulation of tagged antibody, after an adequate time period, at the vicinity of cancer cells, then the penetration of those cells by the tagged antibody without significantly affecting normal cells; (iii) the detection of a signal using an appropriate signal detector; and (iv) the conversion of the detected signal to an image of the cancerous cells.

Other characteristics and advantages of the invention are discussed after this description accompanied by the examples below. In the examples, we will refer to the following figures.

<u>Figure 1: Expression de la protein ICBP90 in HeLa cells (tumour cells) and in pulmonary fibroblasts in primary culture (non-tumoral cells).</u>

The detection of the endogenous protein, ICBP90, was carried out on total protein extracts from confluent (lane 1) and proliferating (lane 2) HeLa

cells and on total protein extracts from primary cultures of human pulmonary fibroblasts at confluence (lane 3) and in proliferation (lane 4). After migration in a polyacrylamide gel in the presence of 8% SDS, the proteins were transferred to nitrocellulose membranes by electrotransfer. The revelation of the protein was performed using antibody 1RC1C-10 diluted to 1/4000 (initial concentration 2 mg/ml) and a secondary antibody coupled to alkaline phosphatase and directed against the heavy chains of mouse antibodies. In the lanes corresponding to extracts from HeLa cells, there is a major band at 97 kDa; for proliferating HeLa cells, supplementary bands of sizes less than 97 kDa appear (lane 2). In confluent human pulmonary fibroblasts, the endogenous protein is not expressed (lane 3), while the protein does appear when the cells begin to proliferate (lane 4). These observations suggest that the endogenous ICBP90 protein is a marker of cellular proliferation for normal cells (fibroblasts), whereas for tumour cells, it is a marker regardless of the cellular stage.

<u>Figure 2: Immunoprecipitation of the endogenous protein</u>

Immunoprecipitation was carried out on total protein extracts from MOLT-4 cells. 1RC1C-10 antibodies were attached to the protein beads of G-Sepharose, then put into contact with protein extract for 2 hours at room temperature. After washing, the bead/1RC1C-10/protein complexes were precipitated by centrifugation and analysed by migration in a 8% polyacrylamide gel in the presence of SDS. They were then transferred to nitrocellulose membranes for revelation of the proteins as indicated in figure 1. A

unique band appears at 97 kDa, as well as a band of 45 kDa corresponding to the heavy chain of 1RC1C-10.

Figure 3: Nuclear localisation of the endogenous protein

We used HeLa cells to examine the endogenous expression of the protein ICBP90 *in situ* employing 1RC1C-10 antibody and a secondary anti-mouse antibody coupled to fluorochrome CY3. The fluorescent marker localises exclusively in the nucleus. The nucleolus and the cytoplasm are not labelled.

Figure 4: Expression of endogenous ICBP59 in proliferating cells

We observed endogenous protein in paraffin sections of human appendix. After deparaffinization and pre-treatment by heat in acid buffer (unmasking of antigenic sites), the sections were incubated for 16 hours with 1RC1C-10 antibodies diluted 1/10000 (initial concentration of 2 mg/ml). Revelation was performed by adding biotinylated secondary antibody, and then incubating with streptavidine-peroxidase complex. A counter-staining of nuclei by Harris' haematoxylin was also carried out. The labelling by 1RC1C-10 is localised essentially in zones of cellular proliferation. The labelled cells are found in glandular crypts (GC), as well as germinative zones (ger).

Figure 5: Expression of ICBP-59 in diverse human tissues

We evaluated the level of expression of mRNA corresponding to ICBP59 in 50 different human tissues

using an RNA dot blot. The blot was hybridised for 16 hours at 68°C with a cDNA (32P) radioactive probe of 679 bp in ExpressHyb (Clontech) hybridisation solution. After washing several times, we revealed the protein by autoradiography (one week exposure at 80°C). The tissues demonstrating the highest expression level were foetal and adult thymus, as well as adult bone marrow and foetal liver.

Figure 6: Nucleotide sequence of ICBP90

cDNA coding for ICBP90 measures 2379 bp. The portions of sequence indicated in bold are those that do not appear in the human EST database (human dbest). The other sequences exist in diverse EST:
From 1 to 325: EST n° AI083773,
From 367 to 865 EST n° AA811055.
From 940 to 1857 EST n° AA488755, EST n° AA129794 and EST n°AA354253.

Figure 7: Protein sequence of ICBP90

The amino acid sequence of ICBP90 was deduced by translation of the nucleotide sequence from figure 6. ICBP90 is composed of 793 residues and has a theoretical molecular weight of 89,758 kDa. The pKi is 7.7. The amino acids indicated in grey correspond to ICBP-59.

Figure 8: Detection of ICBP90 in the sera of patients displaying elevated serum markers for solid tumours.

A volume of 2 µl of serum from each patient was diluted in 1 ml of PBS (1X Phosphate Buffered Saline) containing 0.1% Tween-20 followed by serial dilutions carried out in the same buffer as indicated in the

figure. A 0.5 ml aliquot of each dilution was filtered onto a nitrocellulose membrane using a "Slot Blot BioRad" apparatus. The membrane was then blocked in the presence of PBS buffer (containing 0.1% Tween-20 and 5% milk) for 1 hour at room temperature. The protein ICBP90 was revealed by 1RC-1C10 antibodies (1 ng/ml) and anti-mouse secondary antibodies coupled to peroxidase diluted by 1/5000. The bands were uncovered by chemiluminescence (10 second exposure of X-MAT (Kodak) film).

Figure 9: Structural organisation of the ICBP90 gene.

A. Exons are represented by the boxes: the grey boxes represent coding exons; white boxes represent non-coding exons. The size of each exon is indicated in bp in each box, and the names of the exons are above the boxes. Introns are illustrated schematically by fine lines and their approximate sizes are in bp. A putative transcription start site and a polyadenylation consensus signal are indicated. The ATG is the start codon marking the beginning of translation and TGA, the stop codon for the end of translation.

B. Sequence of the 5′ flanking region of the ICBP90 gene (Seq ID N° 12) (Genbank Accession N° AF 220 226 submitted 30 December 1999). The exons are uppercase and the introns are lowercase. The start codon ATG is in bold uppercase, the boxes rich in GC (GC) and the CCAAT (CB) boxes are in bold lowercase.

Figure 10: Analysis of the ICBP90 promoter.

The promoter sequence of ICBP90 was ligated to the reporter gene, CAT, contained on the pBLCAT2 vector and subsequently transfected into COS-1 cells.

A schematic representation of the constructions appears on the left, the numbers referring to the nucleotides upstream of the start codon. Relative CAT activity of cellular extracts compared to induction of CAT activity by the minimal TK promoter are expressed in percentage (based on the results of 3 independent transfection experiments) and indicated on the right.

Figure 11: Northern and Western blot analysis of the expression of ICBP90.

A. Northern hybridisation was performed on a Northern blotting membrane, containing samples of RNA from cell lines of cancers and various organs. A specific probe for ICBP90, synthesized by PCR, and labelled by digoxigenin, was used to detect ICBP90 mRNA. mRNA sizes are noted on the right side of line 7.

Lines 1 through 7 represent RNA from, respectively, leukaemic HL-60 promyelocytes, hela 53 cells K562 cells from chronic myelogenic leukaemia, MOLT-4 lymphoblastic leukaemia cells, Raji cells from Burkitt's lymphoma, SW480 cells from colorectal adenocarcinoma, and A549 cells from pulmonary carcinoma.

The histogram demonstrates the rate of expression of mRNA corresponding to 5.1 kb and 4.3 kb bands by percentage of the rate of mRNA expression of the 5.1 kb band of HL-60 (line 1, figure 11A).

B. Western blot analysis of ICBP90 expression in MOLT-4 and HeLa cells.

We prepared total cell lysates from proliferating HeLa and MOLT-4 cell cultures. The expression of ICBP90 was analysed by Western blotting using 1RC1C-10 antibodies.

EXAMPLE 1: EVIDENCE OF A NEW BINDING PROTEIN FOR THE ICB SEQUENCE

1.1 Reporter construction for the screening of the library

The simple hybrid system is a powerful technique for detecting, *in vivo*, in yeast the interaction of proteins with specific DNA sequences when screening cDNA libraries. This technique allows you to evaluate directly cDNA corresponding to the protein to be linked. Several studies using this technique resulted in the identification of novel proteins. The protocols are well described by Inouye *et al.* (1994) and Wang and Reed (1993).

Briefly, the following oligonucleotides have been synthesized: 5′-AATTCG**ATTGG**TTCTG**ATTGG**TTCTG**ATTGG**TTCTT-3′ and 5′-CTAGAAGAA**CCAAT**CAGAA**CCAAT**CAGAA**CCAAT**CG-3′. These nucleotides were then hybridised. According to the documentation of the manufacturer (Clontech, Palo Alto, CA), the reporter construct targeted possesses three copies in tandem of the ICB2 sequence (ICB2X3). As mentioned above, one copy of ICB2 is underscored and the CCAAT sequences are in bold. To determine the specificity of protein binding to the ICB box, the following oligonucleotides, containing three copies in tandem of the GC1 box (GC1X3), also present in the promoter, have been synthesized and hybridised:

5′-AATTC**GGGGCGGG**GCC**GGGGCGGG**CCC**GGGGCGGG**GCT-3′

5′-CTAGAG**CCCCGCCC**CGG**CCCCGCCC**CGG**CCCCGCCC**CGG-3′

The resulting target DNA fragments were cloned into the polylinker of the pHISi-1 integrative plasmid (Clontech) by cohesive-end ligation to the plasmid's

XbaI-EcoRI site, upstream of the minimal promoter of the gene, *his3*. The yeast strain, YM4271 (Clontech), was used for the transformation. Transformed colonies of yeast containing the plasmid integrated in their genomes were selected by cultivating the yeast in synthetic dropout medium lacking histidine. We isolated two colonies: one for ICB2 and the other for the GC1 box.

1.2 Screening the library

A cDNA library from the Jurkat cell line, cloned into the EcoRI site of the polylinker downstream of GAL4-AD of the pGAD10 vector (Clontech), was used for screening according to the manufacturer's instructions. Positive clones were selected, and then cultivated in selective medium depleted of histidine and leucine. The plasmid DNA of the clones was recuperated and introduced by electroporation into the bacterial *Escherichia coli* strain, XL1-blue. The sequencing of the inserts were carried out on a matrix of plasmid DNA purified from a 1.5 ml culture using a mini preparation kit (Bio-Rad, Hercules, CA, USA). A cDNA library of human thymus cloned in λgt10 (Clontech) was screened by plaque hybridisation to recuperate a cDNA coding for the N-terminal part of the protein.

1.3 Discovery of ICBP-59

The cDNA from four clones selected using the simple hybrid system was sequenced, then analysed employing a digital database (Genbank, EMBL, PDB, Swissprot) to determine the nature of the coded proteins. Two of the clones correspond to ribosomal proteins (hRS12 and hRS4), one to a serine-threonine kinase (STPLK-1), and the fourth to a human protein having theoretical molecular weight of 59 kDa

(calculated from the translated sequence) that does not appear in the database.

The cDNA coding for hRS4, hRS12, and ICBP-59, and obtained by EcoRI digestion of positive clones in the pGAD10 vector, were cloned into the EcoRI site of the expression vector pGEX-4T-1 (Pharmacia). The recombinant DNA was then transformed in an adapted mouse *Escherichia coli* strain (BL21). We then used a 500 ml culture of a selected clone once the culture reached a density of 0.5. The overexpression of proteins under study was induced by incubation with IPTG (1 mM) for 2 hours at 37°C. The pGEX-4T-1 vector makes possible the recovery of large quantities of proteins fused to glutathione S-transferase (GST). The GST fusion proteins are then purified using Sepharose beads coupled to glutathione (Pharmacia) followed by overnight cleavage with thrombin (0.05 U/ml) at 4° C (Pharmacia).

To test the ability of the 59 kDa protein to bind specifically to the ICB1 and/or ICB2 boxes, three tandem copies of ICB2 (ICB2X3, sequences described above) were labelled at the terminal end with 32 P phosphore using the T4 polynucleotide kinase (New England Biolabs) and [Λ^{32}P]ATP (160 mCi/mmol, ICN Irvine, CA, USA). To examine the specificity of the binding, oligonucleotides containing only one copy of the CCAAT box were synthesized:

ICB1: 5'-AGTCAGGG**ATTGG**CTGGTCTG-';
5'-CAGACCAG**CCAAT**CCCTGACT-3'

ICB2: 5'-AAGCTACG**ATTGG**TTCTTCTG-3';
5'-CAGAAGAA**CCAAT**CGTAGCTT-3'.

The ICBP-59 protein (1 μg) was incubated with 1 ng of oligonucleotide and labelled at its terminal end by

phosphorous ^{32}P in 12% glycerol, 12 mM HEPES-NaOH (pH 7.9), 60 mM KCl, 4 mM Tris-HCl (pH 7.9), 100 ng BSA, 0.6 mM DTT, and 100 ng poly(dI/dC) in 20 µl (Inouye *et al.*, 1994). After a 30-minutes incubation at room temperature, the reaction mix was loaded in 6% polyacrylamide gels. In competition experiments, the quantity indicated of non-labelled oligonucleotides were added to the reaction mix 10 minutes before the addition of proteins. To examine the binding properties of ICBP90 with regard to the ICB2 box, we used the same protocol except that labelled oligonucleotide contained only one copy of the CCAAT sequence as described below:

ICB2: 5'-ATAAAGGCAAGCTACG**ATTGG**TTCTTCTGGACGGAGAC-3'

5'-GTCTCCGTCCAGAAGAA**CCAAT**CGTAGCTTGCCTTTTAT-3'

Binding specificity was studied using a non-labelled nucleotide containing a GC box of the human topoisomerase IIa promoter:

5'-GAATTCGAGGGTAAA**GGGGCGGGG**TTGAGGCAGATGCCA-3'

5'-TGGCATCTGCCTCAA**CCCCGCCCC**TTTACCCTCGAATTC-3'.

These gel retardation experiments in acrylamide gels has given us evidence that the new 59 kDa human protein can specifically bind an ICB DNA sequence. We have called this protein ICBP-59 (for inverted CCAAT Box Binding Protein of 59 kDa).

EXAMPLE 2: CHARACTERISATION OF THE ICBP90 PROTEIN

2.1. Synthesis of antibodies

Mouse monoclonal antibodies were synthesized in our laboratory by injection of ICBP-59 protein using traditional methods (Brou *et al.*, 1993); the protein was purified beforehand by a fusion GST system. Two monoclonal antibodies from 1RC1C-10 and 1RC1H-12 were selected for their ability to detect the ICBP-59

endogenous protein; their specificity was demonstrated in both Western blotting and immunocytochemistry experiments. Before use, the antibodies were purified on a DEAE-cellulose column (DE52, Whatmann) from ascites fluid.

2.2 Detection of the endogenous protein by Western blotting

To detect endogenous ICBP-59 protein, we first used 1RC1C-10 in a Western blot (0.4 μg/ml 1RC1C-10 monoclonal antibodies) of nuclear extracts from confluent and proliferating HeLa cells (Figure 1). COS-1 and HeLa cells were cultivated as previously described (Brou *et al.*, 1993; Gaub *et al.*, 1998; Rochette-Egly *et al.*, 1997). MOLT-4 cells were cultured in 100% air in RPMI supplemented with 10% foetal calf serum. Primary cultures of human pulmonary fibroblasts were prepared and grown in DMEM/F12 as previously described (Kassel *et al.*, 1998). We purchased nuclear extracts of Jurkat cells from Sigma, while we prepared the extracts from MOLT-4 and HL60 as already described in the literature (Lavie *et al.*1999). Proliferating HeLa cells and human pulmonary fibroblasts were obtained by depleting their culture media of serum for 30 hours, then reintroducing foetal calf serum to a concentration of 10% (v/v) for 16 hours. Proliferation was arrested when the cells reached 60 to 70% confluence. Cell cultures stopped at confluence (100% confluence) were prepared in the same way, omitting the serum depletion step. For these two types of cells, total cellular extracts were prepared by first harvesting the cells in PBS (phosphate buffered saline), then sonicating them. Immunotransfer

experiments on total cell lysates and nuclear extracts involved loading the material on 8% SDS polyacrylamide gels and performing a one-dimensional electrophoresis. The proteins were transferred to nitrocellulose membranes that had been blocked with 10% blocking reagent (Roche Molecular Biochemical, Mannheim, Germany). They were then incubated with 1RC1C-10 purified monoclonal antibodies at a concentration of 0.5 µg/ml. A sheep anti-mouse antibody coupled to alkaline phosphatase (fragments Fab, Roche Molecular Biochemicals) was used at a 1/2500 dilution. The signals were detected using 4-nitro blue tetrazolium 5-bromo-4-chloro-3-indolyl-phosphate chloride as substrate.

These experiments show that the endogenous protein has a molecular weight of approximately 97 kDa. Moreover, we observed that the form of the protein varies according to its tumoural or non-tumoural nature, as well as the state of confluence or proliferation of the cells. For example, in the lanes corresponding to extracts from HeLa cells, there is a major band at 97 kDa; for proliferating HeLa cells, supplementary bands of sizes inferior to 97 kDa appear (lane 2). In confluent human pulmonary fibroblasts, the endogenous protein is not expressed and appears when the cells begin to proliferate (lane 4). These observations suggest that the endogenous protein ICBP90 is a marker of cellular proliferation in normal cells (fibroblasts), while, in tumour cells, it would be a marker at any cellular stage.

The use of monoclonal antibodies in immunoprecipitation experiments on nuclear protein extracts, followed by Western blotting, further puts in

evidence the presence of a 97 kDa protein (Figure 2).

The results obtained from Western blotting, for both nuclear protein extracts and immunoprecipitations, show that the 59 kDa protein isolated by the simple hybrid system constitutes a fragment of the corresponding human endogenous protein, in this case, the C-terminal fragment from residue D263. It was, therefore, necessary for us to undertake a new screening of the cDNA library.

2.3. Multiple Human Tissues RNA Dot Blot Analysis

In order to choose a librairy providing us with the best possible chance to isolate the complete protein, we wanted to identify a human tissue expressing the corresponding messenger RNA (mRNA) With a 32P labelled cDNA probe covering part of the ICBP59 sequence, we tested the mRNA expression of interest in 50 different human tissues against a RNA Dot Blot. Briefly, a 678 base pair probe corresponding to the ICBP90 amino acids sequence 269 to 500 was synthesized by PCR using Taq polymerase (Sigma, St Louis, MO, USA). The probe labelled by random priming using dCTP - α 32P was purified on Sephadex G50 columns (Pharmacie, Uppsala, Sweden).

A multiple organ RNA Dot Blot containing poly(A); RNA from 50 different human tissues was hybridised for 20 hours under strong stringency conditions in an ExpressHyb environment (Clontech) at 68° C with a 32P labelled probe. High stringency washing was completed in 0.1 x SSC, 0.1% SDS at 68° C (De Vries et al., 1996).

The results obtained (fig. 5) show that tissues expressing most strongly the ICBP-59 protein mRNA are adult and foetal thymus, as well as adult bone marrow

and foetal liver. Therefore, to isolate the whole protein, we choose an adult thymus cDNA librairy.

2.4. Librairy Screening and ICBP90 Cloning

The bank screening permitted us to obtain several clones of about 4000 base pairs (bp) containing a 2379 bp open reading frame (Fig. 6). This sequence codes for a 793 amino acid protein (Fig.7), which theoretical molecular weight (calculated from the translated sequence) is 89.758 kDa. We called this protein ICBP90 (for Inverted CCAAT Box Binding Protein of 90 kDa) by analogy to the initial 59 kDa protein name.

The ICBP90 cDNA (2379 bp) was synthesized by PCR using Deep Vent DNA polymerase (New England Biolabs, Beverly, MA, USA) and oligonucleotides used during this PCR reaction were near the EcoRI site. The product of the reaction was thereafter sub-cloned in a pGEX-4T-1 vector (Pharmacie) for the GST fusion protein expression in BL21. The over expression was induced by IPTG (1mM) for 4h at 25°C. The ICBP90 protein was then purified.

2.5. Immunocytochemistry and Immunohistochemistry.

The direct observation of the ICBP90 protein on cells and tissues was also carried out.

COS-1 cells were transfected as describes previously (Brou *et al.*, 1993; Gaub *et al.*, 1998) with the pSG5 vector (Stratagene, La Jolla, CA) in which the ICBP90 cDNA (2379 bp) was sub-cloned in the EcoRI restriction site. The cDNA was synthesized by polymerisation chain reaction (PCR) using Deep Vent polymerase (New England Biolabs) and the oligonucleotides flanking the EcoRI restriction site. Plasmidic construction was verified by sequencing. The immunolabelling of the transfected lleLas and COS-1

cells was achieved as described previously (Brou *et al.*, 1993) with 1RC1C-10 and 1 RC1H-12 monoclonal antibodies, respectively. An indirect labelling with ICBP90 immunoperoxidase and IIα topoisomerase was achieved as described previously (Rio *et al.*, 1987, Devys *et al.*, 1993). Human appendices were embedded in paraffin and fixed in 10% buffered formalin (Sigma). Serial sections (3 gm) were incubated overnight at room temperature with 1 RC1C-10 antibody and with IIα anti-topoisomerase antibody (NeoMarkers, Union City, CA, USA). Antibodies bound in a specific manner are visualized through a complex using streptavidine biotin (LAB/LSAB method, Dako LSAB2 System kit; DAKO, Carpinteria, CA, USA).

In immunocytochemistry the 1RC1C-10 antibody labels the HeLa cells nucleus whereas the nucleolus and the whole cytoplasm are not labelled (Fig. 3). In immunohistochemistry, paraffin-embedded human appendix sections show labelling essentially localized in cellular proliferation zones (Fig. 4). Indeed, the labelled cells were located in the glandular crypts (CG) as well as in the germinative zones (Ger). An identical labelling is obtained when using an IIα anti-topoisomerase antibody, an enzyme essentially expressed in proliferating cells (results non illustrated).

2.6. BLAST Research and Domain Prediction

Studies about on-line BLAST have been carried out based on information from the National Centre for Biotechnology Information at the National Institute of Health (Bethesda, MD, USA). SCANPROSITE and PROFILESCANS were used for protein analysis (Infobiogen, Villejuif, France).

ICBP90 includes a "ubiquitin like" domain in its

first 80 amino acids, two sites of potential nuclear localizations in its C terminal and two zinc finger-like domains, one of which could be implicated in the DNA linkage and the other in protein-protein interactions. Several potential phosphorylation sites by protein kinase C, the casein kinase II, as well as by a tyrosine kinase, were also present.

ICBP90 production and purification using the GST fusion system (same procedure as for ICBP-59) permitted to finally test the complete protein ability to link the ICB type DNA sequences. Its behaviour is identical from top to bottom to that observed for ICBP-59.

Finally, we isolated a new human protein that we called ICBP90 for the reasons evoked above. Its theoretical molecular weight is 89.758 kDa and its apparent molecular weight on acrylamide gel is 97 kDa. This protein is not only localized exclusively in human cell nuclei, but it also presents the ability to bind specifically DNA sequences, in this case CCAAT type sequences. For these reasons, we think that ICBP90 has the possibility to modulate the expression of genes which promoter is provided with CCAAT boxes, possibly in reversed position (ICB). The gene of the human topoisomerase IIa we are especially interested in, and which includes five ICB sequences in its promoter, seems to be one of ICBP90 privileged targets.

These experiences allowed to bring to light the 1RC1C-10 antibody remarkable features, which only labels proliferating cells in the case of non cancerous cells; it labels both proliferating and quiescent cancerous cells; it is usable with 4 different techniques (Western blotting, Immunocytochemistry, immunohistology, immunoprecipitation); it has a very

good affinity and allow for the use of 1/150,000 dilution in immunohistochemistry (13 ng/ml); finally, its use generates nearly no background noise.

Future applications of 1RC1C-10 are primarily for diagnostic and basic research. For anatomo-pathologic diagnostics for instance, it would be quite possible to assess the proliferative state of a given cancerous tissue. Regarding basic research, investigations are in progress in our laboratory in order to determine the exact contribution of ICBP90 to proliferation mechanisms in normal and cancerous cells. However, the use of antibodies will be required to study ICBP90 expression as a function of the cellular cycle, of its precise nuclear localization and of its interaction with other cellular proteins.

At the moment we haven't study the expression of ICBP90 with regards to cellular cycle. Nevertheless, in the case where cancerous cell lineages are confluent or when they are not proliferating, we can detect significant differences of ICBP90 expression (Fig. 1) at least with regard to the 97 kDa form. On the other hand, in the non-cancerous confluent cells (human bronchial smooth muscular cells) the ICBP90 expression is hard to detect (results not illustrated). This was confirmed with histological sections where no quiescent cells were labelled by the antibody. It is therefore possible that ICBP90 is expressed whatever the cellular cycle phase in cancerous cells whereas its expression would vary according to each phase in non-cancerous cells. Therefore, this makes the use of the antibody extremely interesting, as, contrary to other cellular proliferation label such as Ki-67, topoisomerase IIα, cycline E and cycline B1, we would have at our

disposition a label for cancerous tissue proliferating cells that would not depend on the cellular cycle phase. Indeed, the end of the S phase is characterized by a very weak Ki-67expression, cycline E labels cells at the end of phase G1 up to the middle of phase S, and cycline B1 labels cells in phase G2/M (for a review, see Darzynkiewicz *et al.*, 1994). Moreover, it has been shown that PCNA (Proliferating Cell Nuclear Antigen) overestimates the number of proliferating cells in some types of tissues (Roskell and Biddolph, 1999).

ICBP90 plays an important role in cellular proliferation by regulating the expression of genes such as those for topoisomerase IIα. Different strategies aiming at blocking the action of this protein must allow modifying cellular proliferation. Anyway, the uses of the 1RC1C-10 antibody as well as of peptides mimicking the ADN/ICBP90 interaction without generating subsequent physiological effect constitute an interesting possibility. The design of its peptides would be directly inspired from the ICBP90 protein sequence we described. A truncated form corresponding to ICBP59 could be one of the first candidates, for instance.

The simple blockage of ICBP90 expression in order to completely eliminate its influence on genes and, by extension, on cellular proliferation can be considered; it could be carried out either by a classic approach such as obtaining inhibitors of the protein, or by a more modern approach corresponding to the interference technique with-double strand RNA (RNA interference or RNAi) as describes recently by Kennerdell & Carthew (1998).

EXAMPLE 3: ISOLATION AND CHARACTERIZATION OF GENE ICBP90

3.1. Material and Methods

3.1.1. Construction and Screening of a Human Placental Genomic Librairy

After partial digestion with MboI enzyme, the placental genomic DNA was split up according to size on a 10 to 40% sucrose gradient. Fifteen kb DNA Fragments were ligated in a λGEM 12 vector previously digested with BamHI (Promega, Madison Wl, USA). After packaging, phage λ particles were assayed on TAP 90 cells. The genomic librairy contains 3.10^6 plaque-forming units (pfu). 10^6 clones were spread out for analysis. A 620 bp probe corresponding to a 5' terminal extremity of the ICBP90 cDNA used for the screening was labelled with α^{32}P-dCTP by a random priming method (Sambrook *et al.*, 1989). The labelled probe is used according to a classic on plaque hybridisation protocol to screen the genomic librairy (Sambrook *et al.*, 1989). Hybridisation was achieved at 68°C in 5X SSC (15 mM NaCl, 1.5 mM sodium citrate pH 7.0), 5 X Denhardt solution, 100 µg / ml of salmon sperm DNA, and 0.1% SDS, followed by 30 minutes washing in 2X SSC, 0.1% SDS at room temperature.

Two screening steps were completed to purify a positive clone. The positive clone was then digested with NotI enzyme and two fragments of 6 and 10 kb were sub-cloned in pBluescript SK$^+$ vector (Stratagene, La Jolla CA, USA) following a standard protocol (Sambrook *et al.*, 1989).

3.1.2. librairy Screening of Human Thymus cDNA

A bank λGT10 of human thymus cDNA 5' end (Clontech, Palo Alto, CA, USA) has been screened by on

plaque hybridisation using the 679 bp cDNA probe synthesized as in the paragraph concerning Northern Blotting Analysis. Signals were detected using 4-nitro-blue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl-phosphate as substratum.

3.1.3. Polymerisation Chain Reaction (PCR) on Placental Genomic DNA

Placental genomic DNA was prepared according to a conventional method (Sambrook *et al.*, 1989). For the 5' region of gene ICBP90, inventors used the PCR Advantage©GC genomic kit from Clontech which is adapted to the genomic DNA regions rich in GC. To cover the 3'-flanking regions, Taq polymerase (Sigma, St Louis, MO, USA) and its corresponding buffer was used. Reactions were achieved according to the manufacturer's instructions while using 250 ng of placental genomic DNA as matrix in a final volume of 50 µl. In order to obtain the 19 kb and 8.7 kb long intron amplification, the PCR Expand™ 20kb^{plus} system (Roche Diagnostics, Mannheim, Germany) was used.

The reaction was completed in 100 µl using 125 ng of placental genomic DNA by reaction.

3.1.4. Plasmidic Constructions and CAT Assays

A set of various fragments was obtained by PCR in the 5' flanking region of gene ICBP90 using 20 nucleotide primers in order to obtain the construction described in fig. 10. These contain a BamH1 restriction site, and a human placental genomic DNA was used as primer. The PCR products were digested and sub-classified upstream from the chloramphenicol acetyl transferase (CAT) reporter gene of a vector containing the thymidine kinase minimal promoter (pBlCAT2). Plasmidic constructions were verified by sequencing.

COS-1 cells were cultivated in a Dulbecco milieu modified by Eagle (DMEM) supplemented with 5% foetal calf serum. After the spreading, the cells were transferred with the various plasmidic constructions (5 µg) using the co-precipitation technique with calcium phosphate (Banerji *et al.*, 1981). Analyses of CAT expression were then carried out as describes elsewhere (Goetz *et al.*, 1996)

3.1.5. Chromosomal Localization of Gene ICBP90

Some metaphasic chromosomes were prepared from human peripheral blood leukocytes according to standard protocols (Haddad et al., 1988). Briefly, a 10 kb probe corresponding to a 5' terminal fragment of the 16 kb clone isolated from the placental genomic DNA screening librairy, was labelled with biotine-16-dUTP (Roche Diagnostics) by "nick-translation". The probe was then precipitated with an excess (50X) of Cot-1 human DNA (Life Technologies, Rockville MD), resuspended in 50% formamide, 1X SSC, pre-hybridised for 2 hours at 37°C then hybridised overnight at 37°C. The detection was carried out using avidin-FITC (Vector Laboratories, Burlingam CA). Chromosomes were counter-stained with 4'-6-diamino-2-phénylindole (Sigma).

3.1.6. Northern and Western Blotting Analysis

A Northern Blotting membrane containing 2 µg of polyA+ RNA by line, coming from 7 different human cancerous cell lines (Clontech) was pre-hybridised in Express Hyb (Clontech), then hybridised with the specific ICBP90 probe in Express Hyb at 68°C for two hours. The double-strand probe labelled with digoxigenin was prepared from PCR amplification of a 676 bp fragment from ICBP90 cDNA (nucleotides 806 to 1485; Genbank accession number AF 129507) according to

the manufacturer's instructions (Roche Diagnostics).

After purification through a Micro Bio-Spin 30 chromatography column (Bio Rad, Hercules, CA), the specific ICBP90 probe (5 ng/ml) was heated at 95°C for 15 minutes then cooled on ice before addition of the hybridisation solution. Washing after hybridisation were carried out twice in 2X SSC, 0.1% SD (30 minutes per wash at room temperature), then twice in SSC 0.1X, 0.1% SD (30 min per wash at 68°C). The membrane was treated with solution A (0.1 M malic acid, 0.15 M NaCl at pH 7.5) then blocked by incubation with 1% blocking agent (Roche Diagnostics) in buffer A for 30 min at room temperature.

An antibody conjugated to alcaline-phosphatase directed against the digoxigenin (Fab fragment, Roche Diagnostics) was added (150 mU/ml) then incubated for 30 min at room temperature. The membrane was then washed twice with solution A, then balanced in 0.1 M tris-HCl, 0.1 M NaCl, pH 9.5. For the detection by chemiluminescence, the inventors used agent disodium 3-(4-methoxyspiro(1,2-dioetane-3,2'-(5'-chloro)tricyclo-[3,3.1.$1^{3,7}$]decan}-4-yl) phenyl phosphate© (Roche Diagnostics) according to the manufacturer's instructions. mRNA strips were quantified using the NIH software Image 1.62 and expressed as a percentage of the most abundant mRNA strip (e.g. the 5.1 kb strip of HL-60 cells).

Western Blotting analysis was carried out as describes elsewhere (Hopfner et al., 2000). Signals were detected using 4-nitro-blue tetrazolium chloride and 5-bromo-4chloro-3-indolyl phosphate as substrate.

3.1.7. Local Base Alignment Research Tools, Primer Transcription and PolyA Signal Sites Predictions

Local base alignment research tools was completed via the National Biotechnology Information Center at the National Institute of Health (Bethesda, MD, USA). The transcription factor librairy screening with Mat Inspector software, the primer transcription site predictions (TSS) with Neural Network, as well as the polyA signal prediction, were all carried out at Baylor College of Medicine (Reese *et al.*, 1996).

3.2. Results

3.2.1. Isolation and Characterization of Gene ICBP90

A DNA complementary librairy of human placenta cloned within the lambda GEM 12 phage was screen using a DNA probe. The screening lead to the purification of a single positive clone with a 16 kb insert. The sequence analysis permitted to determine that it contained a 10 kb intronic sequence containing 3 exons (called B, C, and D in fig. 9A). All others screenings, namely including those completed with PCR on BAC (Bacterial Artificial Chromosome) or YAC (Yeast Artificial Chromosome) banks, failed to yield other positive clones. Therefore, we decided to determine the remainder of the gene organization directly by PCR on human placenta genomic DNA. The biggest difficulty was to get the 5' end of the 19 kb intron. Primers were so chosen in exon A (sense primer) and in the 5′ end of the 16 kb clone (anti-sense primer). The exon E and the 8.7 kb intron were amplified using a sense primer in exon D and anti-sense primer in exon F. Finally, the complete sequence of exon F up to the poly-adenylation signal was determined using a sense primer chosen at the beginning of exon F and the anti-sense primer in the 3' end of an EST (reference in GenBank No.

AW297533) homologous to Gene ICBP90 sequence. The complete sequence of gene ICBP90 shows that it is made of 6 coding exons which size varies from 100 bp to 3453 bp. Most exon/intron junctions match consensus sequences for splicing acceptor and donor sites. A poly-adenylation (AATAAA) consensus sequence was found in the 3' region, e.g. 1152 nucleotides after the stopping codon in fig. 9A.

3.2.2 The 5' Region of Gene 11OEP90

The complementary DNA screening librairy of human thymus cloned in lambda gt 10 phage lead to obtaining two cDNA populations distinguishing one another from their 5' region, precisely 10 base pairs upstream from the primer codon, i.e. in the non-translated 5′ region. These two cDNA populations predict the existence of two alternative exons in 5′ called exon I and II (Fig. 9A). We observed that exons I and II are linked to an alternative internal splicing site of exon A. Moreover, in a database, we found an EST (reference in GenBank No. AI084125) corresponding to nucleotides 1290 to 1356 (Fig. 9B). The positions of these two exons and of the EST inside the locus were determined by PCR. For that, we used primers corresponding to the first 18 nucleotides of each exon and an anti-sense primer selected from the first exon translated (exon A). This strategy permitted us to rebuild the 5' region as represented in fig. 9A and 9B, with exon I corresponding to nucleotides 1 to 134 and exon II corresponding to nucleotides 676 to 725. The EST sequence (AI084125) is adjacent to exon A internal splicing site. We haven't determine yet with precision the beginning of exons I, II, and A since their sequences have been deducted from cDNA bank screenings

(Fig. 9A).

Four GC boxes (GC1 to GC4) have been found in the 5' region (Fig. 9B). These boxes represent the potential sites of linkage for the Spl transcription factor, but only one box (GC3) corresponds to a consensus sequence, e.g. GGGGCGGGG. Besides two CCAAT boxes (CB1 and CB2) were found. Predictive analyses of sequences suggest that two promoter regions exist in the 5' region, e.g. before the initiation codon (ATG). Two potential transcription initiation sites have been predicted in positions 571 and 827. The first follows the linkage consensus sequence of Spl and the second follows the GC1 box (between exons I & II, and exons II & A, respectively). In order to determine if these two regions are functional as promoter region, several plasmidic constructions containing a reporter gene (the Chloramphenicol Acetyl Transferase gene; CAT) downstream from the various potential promoters regions were prepared. COS cells were transfected with these plasmidic constructions. Fig. 10 shows the results obtained corresponding to a percentage of increased basal activity. The maximal activity was obtained with the plasmidic construction containing 1114 bp upstream from the translation initiation site, with a 236.7% increase of basal promoter activity (thymidine kinase gene minimal promoter). The plasmidic construction containing 642 bp upstream from ATG lead to a 115.6% increase whereas plasmidic construction containing the sequence solely between exon I and II showed a comparatively weak activity with only a 22.8% increase (fig. 10). These results suggest the existence of a promoter region between exons II and A.

3.2.3. Chromosomal Localization of Gene ICBP90

The chromosomal localization of gene ICBP90 was completed by fluorescence *in situ* hybridisation (FISH). Gene ICBP90 is localized on chromosome 19p13.3 in a telomeric region. A research carried out at Genbank showed that a 6Mb region in the chromosomal strip 19p 13.3 of a chromosome 19 (hybrid human / hamster 5HL2 B) specific cosmid bank contains 147 nucleotides coding for ICBP90 amino acids 746 to 793. This sequence has been localized between the STS (sequence tagged site) markers D 19S883 and D 19S325.

3.2.4 ICBP90 Expression in Various Cellular Lineages

ICBP90 participates in the regulation of the gene TopIIα expression (Hopfner *et al.*, 2000). As TopIIα is expressed 3rd differential manner in various tumours and cellular lineages, ICBP90 itself is susceptible to have a complex regulation in term of activity and genic expression.

In a first step towards understanding the mechanisms regulating gene ICBP90 expression, ICBP90 mRNA was analysed in various cellular lineages. ICBP90 mRNA was studied in the HL60 cellular lineage derived from promyelocytic leukaemia (lineage 1), Hela S3 cells (lineage 2), MOLT-4 lymphoblastic leukaemia cells, Raji Burkitt lymphoma cells (lineage 5), SW 480 colorectal adenocarcinoma (lineage 6), A549 lung carcinoma cells (lineage 7) (fig. 11A).

Two 4.3 and 5.1 kb bands of mRNA are observed. The relative amounts of mRNA in the bands vary according to the cell type. The histogram in Figure 11A shows the levels of mRNA in the bands of each of the cell lines, expressed in percentage of the maximum amount of 5.1 kb bands of mRNA observed in the HL-60 cells (line 1,

Figure 11A). In the MOLT-4 cells, only the 4.3 kb band of mRNA is observed, while in the cells from promyelocytic leukaemia the 5.1 kb band is predominant. In the Raji cells of Burkitt's lymphoma, only the 5.1 kb band is detected. Approximately equal amounts of the two types of mRNA are observed in the other cell lines, that is, the Hela, K562, A549, SW580 cells. For the HL-60 cells, nevertheless, the 5.1 kb mRNA is more strongly expressed than the 4.3 kb mRNA. Other analyses have been undertaken on the Hela cells to confirm that the 2 transcripts originate from the transcription of the ICBP90 gene. A cDNA probe of 626 bp labelled with digoxigenin localized immediately upstream of the poly A signal (that is, the exon F) and used as probe for Northern Blotting experiments, has produced the same results, that is, the appearance of two 4.3 kb and 5.1 kb bands of mRNA. This result confirms that the two forms of mRNA are generated from a single gene.

The inventors have also studied the expression of the ICBP90 protein in order to determine if these two isoforms of mRNA are likely to code for two different proteins.

Figure 11B shows the expression profile of ICBP90 in protein extracts of MOLT-4 and Hela cells. While a single band of 97 kDa is observed in the MOLT-4 cells, in the Hela cells, beside the 97-kDa band that is doubled, several other bands with a lower molecular weight are observed. These results suggest that in the MOLT-4 cells, an mRNA codes for a single form of ICBP90. Conversely, in the Hela cells, the two mRNA are likely to lead to the production of different isoforms of ICBP90.

3.3 Comments

The ICBP90 gene is spread over approximately 35.8 kb. Six translated exons and two untranslated exons, and then, seven introns have been identified by the inventors. The two zinc-finger domains of ICBP90 are coded by the same exon (exon F) in contrast to the receptor gene for human estrogens in which each of the presumed zinc fingers of the DNA binding domain of the receptor are coded separately (Ponglikitmongkol *et al.* (1988)). The "ubiquitin-like" domain of ICBP90 is coded by exons A and B while the "leucine zipper" is coded by exon B. Interestingly, only exon F is likely to code for a functional protein because it codes for two nuclear localization signals, the zinc-finger domains and several presumed sites of phosphorylation. Two large 8.7 kb and 19 kb introns have been found.

The ICBP90 gene has been localized in the chromosome region 19p13.3. Several other genes have been localized in this region, for example the Nuclear Factor I/C (also a CCAAT binding transcription factor) (Qain *et al.* (1995)). Interestingly, an atypical translocation t(7;19) in the acute myelomonocytic leukaemia, involving a fragile site at the 19p13.3 locus has been described (Sherer *et al.* (1991)). Also, it has been suggested that the genes involved in the development of pancreatic carcinomas are localized at 19p13.3 and 19q13.1-13.2 (Hoglund *et al.* (1998)). Rearrangements of the 14q32.3 and 19p13.3 bands with a preferential deletion of the short arm of chromosome 1 form non-random chromosome alterations in multiple myeloma and leukaemia of cells of the plasma (Taniwaki *et al.* (1996)). Other genes have been localized in this region; they include a gene involved in adenocarcinoma of the Peutz-Jeghers syndrome (Gruba *et al.* (1998)).

Also, it has been suggested that the presumed tumour suppressor gene for malignant adenoma is localized on D19S216 at the 19p13.3 chromosome band that plays an important role in tumourigenesis of malignant adenoma (Lee *et al.* (1998)).

The analysis of the sequence of the 5′ region of the ICBP90 gene has revealed the existence of several untranslated exons with a promoter region between exons II and A and probably a second weaker promoter localized between exons I and II. The promoter region between exons II and A is a promoter without TATA sequence suggesting that the ICBP90 gene may be a housekeeping gene at least when this promoter is involved. In this sense, it strongly resembles promoter regions of the genes ATFα (Goetz *et al.*, 1996), CRE-BP1 / ATF2 (Nagase *et al.*, 1990) and TopIIα (Hochhauser *et al.*, 1992) which do not contain canonical TATA boxes but several SP-1 binding sites.

The GC and/or CCAAT boxes are likely to be involved in the regulation of the expression of the ICBP90 gene via transcription factors SP-1 and the CCAAT binding proteins. Furthermore, given that the ICBP90 protein is a CCAAT binding protein, ICBP90 is also likely to regulate its own expression.

A data librairy of transcription factors has been screened with the aid of the Mat Inspector computer program from the Baylor College of Medicine and numerous binding sites of transcription factors have been identified in the sequence preceding the ATG codon (Figure 9B).Among these binding sites for the transcription factors it is interesting to note binding sites of the AP-2 transcription factor regulated during the development and which controls the DR-nm23 gene

expression (Martinez *et al.* (1997)), the binding sites of the "zinc-finger" myeloid protein MZF 1 which is involved in the regulation of hematopoiesis (Hromas *et al.*(1996)).

The Northern Blotting analysis has demonstrated that two populations of mRNA exist, 4.3 kb and 5.1 kb. Interestingly, each population presents a cellular specificity. For example, the lymphoblast cells MOLT-4 only express the 4.3 kb mRNA, while in the Raji cells of Burkitt's lymphoma (mature B lymphocytes), only the 5.1 kb transcript is observed. The HL-60 cells express more 5.1 kb mRNA then 4.3 kb mRNA. The HL-60 cells and the Raji cells of Burkitt's lymphoma are more differentiated than the MOLT-4 cells suggesting that the levels of expression of the 5.1 kb transcript relative to that of 4.3 kb may be directly correlated with the state of differentiation of the cells.

Interestingly, an expressed sequence tag (EST, Expressed Sequence Tag) corresponding to the 5' sequence of the exon A has been identified from anaplastic oligodendroglioma (Genbank Accession No. AI 084 125) while an EST corresponding to the inclusion of exon II has been isolated from a mixture of tumours with germinal cells (Genbank Accession No. AI 968 662). The results of the inventors therefore suggest that the regulation of the ICBP90 transcripts is comparable to that which happens with the oestrogen receptor. In fact, six different transcripts coding a common protein, but differing in the untranslated 5' region because of an alternative splicing of upstream exons, have been reported (Flouriot *et al.*, 1998 and Grandien, 1996).

The Western Blotting analysis shows a major band

at 97 kDa in the MOLT-4 cells while several bands are observed in the Hela cells (Figure 11B). This data is in agreement with the existence of several ICBP90 mRNA and/or isoforms of the ICBP90 protein for which the level of expression may be controlled in a cell-specific manner.

Two protein isoforms for the oestrogen receptor have been described (Griffin *et al.*, 1999) which differ from each other by the 41 N-terminal amino acids. The 97 kDa double band observed from the Hela cells (Figure 11B) is therefore likely to represent two isoforms differing by their N-terminal end. To do this, the exon A coding for 47 amino acids is spliced outside the reading frame, and consequently, the protein-coding region begins with exon B. Nevertheless, it is also possible that there are other exons likely to be transcribed in other tissues.

Also, the 8.7 kb intron (that is, between exon D and E) is likely to contain a promoter region which may lead to ICBP90 isoforms with lower molecular weight than those observed in the Hela cells in proliferation (Figure 11B). Interestingly, the tissue specificity of the different mRNA of the oestrogen receptor is determined by different promoters for which the activity appears to be altered in the cell lines of breast cancer (Flouriot *et al.*, 1998).

All these results suggest that the ICBP90 gene and the ICBP90 protein present characteristics common with members of the family of the receptor for retinoic acid, steroids, thyroid hormones where it concerns gene and protein structures.

In fact, the inventors have demonstrated experimentally, by using the double-hybrid technique,

the existence of interactions between the ICBP90 protein and TIP60 (Tat Interactive Protein, 60 kDa). The TIP60 protein has very recently been described as being a coactivator of the nuclear receptor, especially the receptor for the androgens (Brady ME *et al.*, 1999).

Because of this, ICBP90 is capable of playing the role of a nuclear receptor on which an endogenous ligand is bound. Therefore, it is also within the scope of the present invention to use the ICBP90 polypeptide of the invention to isolate, screen, and identify the endogenous ligand. It is also within the scope of the invention to use the ICB90 polypeptide of the invention to isolate, screen identify natural or synthetic, biological or chemical, agonist or antagonist molecules of this natural ligand.

REFERENCES

Austin *et al.* (1993), Biochem. Biophys. Acta, 1172, 283-291

Banerji, J. *et al.* (1981), Cell, 27: 299-308.

Barany, F. (1991), Proc. Natl. Acad. Sci. USA, 88, 189-193.

Boritzki, T.J. *et al.* (1988), Biochem. Pharmacol., 37, 4063-4068.

Brady, M.E. *et al.* (1999), J. Biol Chem., 274: 17599-17604.

Brandt, T.L. *et al.* (1997), J. Biol. Chem., 272, 6278-6284.

Brou, C. *et al.* (1993), EMBO J., 12, 489-499.

Buckholz, R.G. (1993), Curr. Op. Biotechnology, 4, 538-542.

Burg, J.L. *et al.* (1996), Mol. and Cell. Probes, 10, 257-271.

Chu, B.C.F. *et al.* (1986), Nucleic Acids Res., 14, 5591-5603.

Chung, T.D.Y. *et al.* (1989), Proc. Natl. Acad. Sci. USA, 86, 9431-9435.

Darzynkiewicz *et al.* (1994), Methods in Cell Biology, 41, 421-435.

Deffie, A.M. *et al.* (1989), Cancer Res., 49, 58-62.

DeVries, L. *et al.* (1995), Proc. Natl. Acad. Sci. USA, 92, 11916-11920.

Devys *et al.* (1993), Nature Genet., 4, 335-340.

Drake, F.H. *et al.*, Biochemistry, 28, 8154-8160.

Duck, P. *et al.* (1990), Biotechniques, 9, 142-147.

Edwards, C.P. and Aruffo, A. (1993), Curr. Op. Biotechnology, 4, 558-563.

Erlich, H.A. (1989), New York: Stockton Press.

Flouriot *et al.* (1998), Mol. Endocrinol., 12:1239-254.

Fry, A.M. *et al.* (1991), Cancer Res., 51, 6592-6595.

Furth *et al.* (1992), Anal. Biochem., 205: 365.

Gaub, M.P. *et al.* (1998). J. Histochem. Cytochem., 46, 1103-1111.

Goetz, J. *et al.* (1996), J. Biol. Chem., 271: 29589-29598.

Goswami, P.C. *et al.* (1996), Mol. Cell. Biol., 16, 1500-1508.

Grandien (1996), Mol. Cell. Endocrinol., 116: 207-212.

Griffin *et al.* (1999), Mol. Endocrinol., 13: 1571-1587.

Gruba *et al.* (1998), Cancer Res., 58: 5267-5270.

Guatelli, J. C. *et al.* (1990), Proc. Natl. Acad.

Sci. USA, 87, 1874-1878.

Guinee, D.G. *et al.* (1996), Cancer, 78, 729-735.

Haddad *et al.* (1988), Human Genet. 103: 619-625.

Heck, M.M. *et al.* (1988), Proc. Natl. Acad. Sci. USA, 85, 1086-1090.

Herzog, C.E. and Zwelling, L.A. (1997), Biochem. Biophys. Res. Commun., 232, 608-612.

Hochhauser, D. *et al.* (1992), J. Biol. Chem., 267, 18961-18965.

Hoglund *et al.* (1998), Genes Chromosomes Cancer 21:8-16.

Hopfner *et al.* (2000), Cancer Res., 60:121-128.

Hromas *et al.* (1996), Curr. Top. Microbiol. Chem., 211:159-164.

Innis, M.A. *et al.* (1990), Academic Press.

Inouye, C. *et al.* (1994), DNA Cell Biol., 13, 731-742.

Isaacs, R.J. *et al.* Biochem. Biophys. Acta, 1400, 121-137.

Isaacs, R.J. *et al.* (1996), J. Biol. Chem., 271, 16741-16747.

Jenkins, J.R. *et al.* (1992), Nucleic Acids Res., 20, 5587-5592.

Kassel, O. *et al.* (1998), Mol. Pharmacol., 54, 1073-1079.

Kennerdell, J.R. and Carthew, R.W. (1998), Cell, 95, 1017-1026.

Kievitis, T. *et al.* (1975), J. Virol. Methods, 35, 273-286.

Kohler, G. *et al.* (1975), Nature, 256 (5517), 495-497.

Kubo, T. *et al.*, (1995), Cancer Res., 55, 3860-3864.

Kwoh, D.Y. *et al.*, (1989), Proc. Natl. Acad. Sci. USA, 86, 1173-1177.

Landegren, U. *et al.* (1988), Science, 241, 1077-1080.

Lavie, J. *et al.* (1999), J. Biol. Chem., 274, 2308-2314.

Lee *et al.* (1998), Cancer Res., 58: 1140-1143.

Lim, K. *et al.* (1998), Biochem. Mol. Biol. Int., 46, 35-42.

Lizardi, P.M. *et al.* (1988), Bio/technology, 6, 1197-1202.

Lucknow, V.A. (1993), Curr. Op. Biotechnology, 4, 564-572.

Martinez, *et al.* (1997), Cancer Res., 57: 1180-1187.

Matthews, J.A. *et al.* (1988), Anal. Biochem., 169:1-25.

Miele, E.A. *et al.* (1983), J. Mol. Biol., 171, 281-295.

Nagase *et al.* (1990), J. Biol. Chem., 265:17300-17305.

Nitiss, J.L. (1998), Biochem. Biophys. Acta, 1400, 63-81.

Olins, P.O. and Lee, S.C. (1993), Curr. Op. Biotechnology, 4, 520-525.

Pommier, Y. *et al.* (1994), Cancer Invest., 12, 530-542.

Ponglikitmongkoi *et al.* (1988), EMBO J. 7:3385-3388.

Rio, M.C. *et al.* (1987), Proc. Natl. Acad. Sci. USA, 84, 9243-9247.

Qian *et al.* (1995), Genomics, 28:66-73.

Reese *et al.* (1996), Large Scale Sequencing

Specific Neural Networks for Promoter and Splice Recognition. Biocomputing: Proceedings of the 1995 Pacific Symposium. Edited by Lawrence Hunter and Terri E. Wood Scientific Singapore, 1996, January 27, 1996.

Rochette-Egly, C. *et al.* (1997), Cell, 90, 97-107.

Roskell, D.E. and Biddolph, S.C. (1999), Eur. J. Med. Res. 26, 105-106.

Rolfs, A. *et al.* (1991), Berlin: Springer-Verlag.

Sambrook, J. *et al.* (1989), Molecular Cloning: A Laboratory Manual, Sec. Ed. Cold Spring Harbor Lab., Cold Spring Harbor, New York.

Sandri, M.I. *et al.* (1996), Nucleic Acids Res., 24, 4464-4470.

Segev, D. (1992), Kessler C. Springer Verlag, Berlin, New York, 197-205.

Sherer *et al.* (1991), Cancer Genet. Cytogenet., 57: 169-173.

Stone, B.B. *et al.* (1996) Mol. and Cell. Probes, 10:359-370.

Tang *et al.* (1992), Nature, 356:152.

Taniwaki *et al.* (1996), Leuk. Lymphoma, 21:25-30.

Tsai-plugfelder, M. *et al.* (1988), Proc. Natl. Acad. Sci. USA, 85, 7177-7181.

Walker, G.T. *et al.* (1992), Nucleic Acids Res., 20:1691-1696.

Wang, J.C. (1996), Ann. Rev. Biochem., 65, 635-692.

Wang, M.M. and Reed, R.R. (1993), Nature (London), 364, 121-126.

Yamazaki *et al.* (1996), Acta Oncol., 35, 417-423.

CLAIMS

1. Isolated polypeptide called ICBP90 (inverted CCAAT box binding protein 90) with amino acid sequence SEQ ID No. 2.

2. Isolated polypeptide characterized in that it comprises a polypeptide chosen from:

a) a polypeptide with sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6 or SEQ ID No. 8;

b) a polypeptide variant of polypeptide with sequences of amino acids defined in a);

c) a polypeptide homologous with the polypeptide defined in a) or b) and including at least 80% homology, preferably 90% with said polypeptide of a);

d) a fragment of at least 5 consecutive amino acids of a polypeptide defined in a), b) or c);

e) a biologically active fragment of a polypeptide defined in a), b) or c).

3. Polypeptide according to any one of Claims 1 to 2 and characterized in that it comprises of at least one domain for fixation to the DNA selected in the group composed of a "zinc-finger" domain and a "leucine-zipper" domain.

4. Polypeptide according to Claim 3 characterized in that the DNA sequence on which said polypeptide is bound is a CCAAT box, preferably an inverted CCAAT box (Inverted CCAAT box: ICB).

5. Isolated polynucleotide characterized in that it codes for a polypeptide according to Claim 1.

6. Polynucleotide according to Claim 5 with sequence SEQ ID No. 1.

7. Isolated polynucleotide characterized in that it consists of a polynucleotide chosen from:

a) a polynucleotide with sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5 or SEQ ID No. 7 or in which the sequence is that of the RNA corresponding to the sequence SEQ No. 1, SEQ ID No. 3, SEQ ID No. 5 or SEQ ID No. 7;

b) a polynucleotide in which the sequence is complementary to the sequence of a polynucleotide defined in a),

c) a polynucleotide in which the sequence consists of at least 80% homology with a polynucleotide defined in a) or b),

d) a polynucleotide hybridizing under very stringent conditions with a polynucleotide sequence defined in a), b) or c),

e) a fragment of at least 15 consecutive nucleotides, preferably 21 consecutive nucleotides, and preferably 30 consecutive nucleotides of a nucleotide defined in a), b), c) or d) with the exception of human EST AI 084 125, with the exception of the sequence corresponding to the sequence SEQ ID No. 944 published August 5, 1999 in the patent application WO 99 38972 and with the exception of the sequences SEQ ID No. 9, No. 10, No. 11 corresponding, respectively, to the human EST No. AI 083 773, AA 811 055, No. AA 488 755, N° AA 129 794 and No. AA 354 253.

8. Polynucleotide according to Claim 7 characterized in that it is labelled directly or indirectly with a radioactive compound or a nonradioactive compound.

9. Use of a polynucleotide according to Claim 8 as primer for amplification or polymerization of nucleic sequences.

10. Use of a polynucleotide according to Claim 8

as probe for detection of nucleic sequences.

11. Use of a polynucleotide according to Claim 8 as sense or antisense oligonucleotide to control the expression of the corresponding protein product.

12. Recombinant cloning vector of a polynucleotide according to one of Claims 5 to 8 and/or expression of a polypeptide according to one of Claims 1 to 4 characterized in that it contains a polynucleotide according to any one of Claims 5 to 8.

13. Vector according to Claim 12 characterized in that it consists of the parts enabling the expression [and] possibly the secretion of said polypeptide in a host cell.

14. Vector according to any one of Claims 12 to 13 characterized in that the parts enabling the expression of said polypeptide are chosen form:

a) the isolated polynucleotide with sequence SEQ ID No. 12;

b) a polynucleotide in which the sequence is complementary to the polynucleotide sequence defined in a);

c) a polynucleotide in which the sequence consists of at least 80% identity with a polynucleotide defined in a) or in b);

d) a polynucleotide hybridizing under very stringent conditions with a sequence of the polynucleotide defined in a), b) or c).

15. Host cell, characterized in that it is transformed by a vector according to one of Claims 12, 13 and 14.

16. Method for preparation of a recombinant polypeptide characterized in that a host cell is cultured according to Claim 15 under conditions

enabling the expression and possibly the secretion of said recombinant polypeptide and that said recombinant polypeptide is recovered.

17. Recombinant polypeptide obtainable by a method according to Claim 16.

18. Monoclonal or polyclonal antibody and its fragments characterized in that it specifically binds a polypeptide according to one of Claims 1 to 4 and 17.

19. Monoclonal antibody according to Claim 18 specific for the human ICBP90 protein and capable of inhibiting the interaction between ICBP90 and the DNA sequence on which the protein ICBP90 is specifically bound.

20. Monoclonal antibody according to Claim 18 specific for the human ICBP90 protein and capable of inhibiting the interaction between ICBP90 and proteins with which ICBP90 interacts, said proteins preferably being ICBP90 itself or proteins of a transcriptional complex.

21. Method for detection and/or measuring of a polypeptide according to one of Claims 1 to 4 and 17 in a biological sample, characterized in that it comprises the following steps:

a) putting the biological sample in contact with an antibody according to one of Claims 18 to 20;

b) revealing a formed antigen-antibody complex.

22. kit for making use of a method according to Claim 21 in a biological sample by immunological reaction, characterized in that it comprises the following parts:

a) a monoclonal or polyclonal antibody according to one of Claims 18 to 20;

b) if applicable, the reagents for the formation

of the favourable medium for the immunological reaction;

c) the reagents enabling the detection of the antigen-antibody complex produced by the immunological reaction.

23. Method for detection and/or measurement of a polynucleotide according to any one of Claims 5 to 8 in biological sample, characterized in that it consists of the following steps:

a) isolating the DNA from the biological sample to be analyzed, or obtaining cDNA from the RNA of the biological sample;

b) specific amplification of the DNA with the aid of primers according to Claim 9;

c) analysis of amplification products.

24. Kit for making use of a method according to Claim 23 in a biological sample characterized in that it comprises the following parts:

a) a pair of nucleic primers according to Claim 9;

b) the reagents necessary for carrying out an amplification reaction of DNA;

c) possibly a component enabling the verification of the sequence of the amplified fragment, more particularly a probe according to Claim 10.

25. Method for detection and/or measurement of a nucleotide according to any one of Claims 5 to 8 in biological sample characterized in that it consists of the following steps:

a) putting a probe according to Claim 10 in contact with a biological sample;

b) detection and/or measurement of the hybrid formed between said probe and the DNA of the biological sample.

26. Kit for making use of a method according to Claim 25 in a biological sample characterized in that it comprises the following parts:

a) a probe according to Claim 10;

b) the reagents necessary for using a hybridization reaction.

27. Method according to Claims 21, 23 and 25 for the diagnosis of cellular proliferation.

28. Ligand screening method likely to affect the transcription activity of a gene the promoter of which consists of CCAAT and/or inverted CCAAT boxes (ICB) likely to bind a polypeptide according to Claims 1 to 4 and 17 and which consists of the following steps:

a) putting said polypeptide and one or more potential ligand(s) in the presence of reagents necessary for using a transcription reaction;

b) detection and/or measurement of the transcription activity.

29. Ligand screening method likely to affect the "nuclear receptor" function of the polypeptide according to Claims 1 to 4 and 17 and which consists of the following steps:

a) putting said polypeptide and one or more potential ligands in the presence of the necessary reagents;

b) detection and/or measurement of the transcription activity of a gene the promoter of which comprises CCAAT and/or inverted CCAAT (ICB) boxes likely to bind said polypeptide.

30. Kit for making use of a method according to Claims 28 and 29 in a biological sample characterized in that it comprises the following parts:

a) a polypeptide according to Claims 1 to 4 and

17;

b) a ligand;

c) the reagents necessary for using a transcription reaction.

31. Ligand obtainable by the method according to Claim 28 or 29.

32. Compound characterized in that it is chosen from:

a) a polypeptide according to one of Claims 1 to 4 or 17;

b) a polynucleotide according to one of Claims 5 to 8;

c) a vector according to one of Claims 12 to 14;

d) a cell according to Claim 15;

e) an antibody according to one of Claims 18 to 20;

f) a ligand according to Claim 31

as a drug.

33. Compound according to Claim 32 as a drug intended for the prevention and/or treatment of cancer.

34. Use of a compound according to Claims 32 and 33 for the preparation of a drug intended to modulate, increase or decrease cell proliferation.

35. Pharmaceutical composition for the preventive and curative treatment of cancer characterized in that it contains a therapeutically effective amount of a compound according to one of Claims 32 and 33 and a pharmaceutically acceptable vehicle.

36. Pharmaceutical composition characterized in that it comprises an antibody according to one of Claims 18 to 20 as screening agent conjugated with at least one agent selected from the group of antiproliferative, antineoplastic or cytotoxic agents.

37. Product comprising at least one compound according to Claims 32 and 33 and at least another anticancer agent as combination product for simultaneous use, separate use or spread over time in anticancer therapy.

38. Composition for the detection, localization and imagery of cancers, comprising an antibody according to any one of Claims 18 to 20, the antibody is labelled directly or indirectly with a marker generating a signal selected from radioactive isotopes and nonisotope entities.

39. Method for the detection, localization and imagery of cancer, comprising the steps of:

a) parenteral injection of a composition according to Claim 38 in a human being;

b) accumulation after sufficient time of the labelled antibody at the cancer cells, then penetration of the labelled antibody within said cells, without said antibody being bound substantially to the normal cells; and

c) detection of the signal by means of a signal detector; and

d) conversion of the detected signal into an image of cancer cells.

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

(43) Date de la publication internationale
28 décembre 2000 (28.12.2000)

PCT

(10) Numéro de publication internationale
WO 00/78949 A1

(51) Classification internationale des brevets[7]: C12N 15/12, C07K 14/435, 14/47, C12N 15/10, 15/66, 15/11, C12Q 1/68, C07K 16/18, G01N 33/53, A61K 51/00, A61P 35/00

(21) Numéro de la demande internationale: PCT/FR00/01747

(22) Date de dépôt international: 22 juin 2000 (22.06.2000)

(25) Langue de dépôt: français

(26) Langue de publication: français

(30) Données relatives à la priorité:
99/07935 22 juin 1999 (22.06.1999) FR

(71) Déposant *(pour tous les États désignés sauf US)*: ASSOCIATION POUR LE DEVELOPPEMENT DE LA RECHERCHE EN GENETIQUE MOLECULAIRE (ADEREGEM) [FR/FR]; 231, rue de Charenton, F-75012 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants *(pour US seulement)*: BRONNER, Christian [FR/FR]; 19, rue Exelmans, F-67640 Fegersheim (FR). HOPFNER, Raphaël [FR/FR]; 4, rue Fix, F-67000 Strasbourg (FR). MOUSLI, Marc [FR/FR]; 10, rue de Libreville, F-67400 Illkirch (FR). JELTSCH, Jean-Marc [FR/FR]; 1, rue de la Bruche, F-67120 Molsheim (FR). LUTZ, Yves [FR/FR]; 12, rue d'Ypres, F-67000 Strasbourg (FR). OUDET, Pierre [FR/FR]; 17, rue Vauban, F-67000 Strasbourg (FR).

(74) Mandataires: MARTIN, Jean-Jacques etc.; Cabinet Regimbeau, 26, avenue Kléber, F-75116 Paris (FR).

(81) États désignés *(national)*: AU, CA, JP, US.

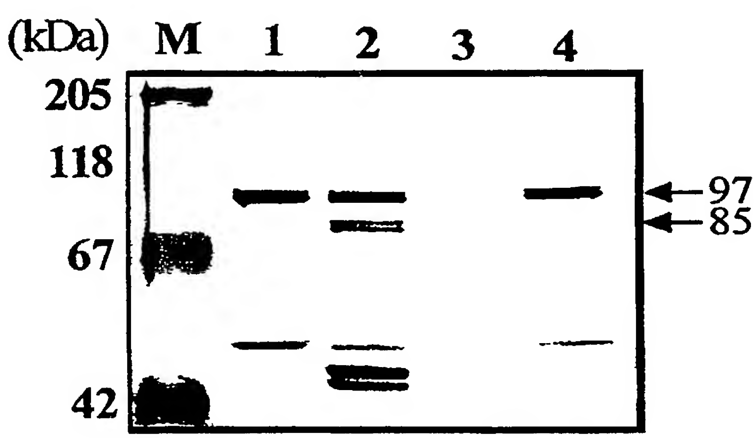
(84) États désignés *(régional)*: brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Publiée:
— *Avec rapport de recherche internationale.*

[Suite sur la page suivante]

(54) Title: ICBP90 POLYPEPTIDE AND ITS FRAGMENTS AND POLYNUCLEOTIDES CODING FOR SAID POLYPEPTIDES AND APPLICATIONS FOR DIAGNOSING AND TREATING CANCER

(54) Titre: POLYPEPTIDE ICBP90 ET SES FRAGMENTS ET POLYNUCLEOTIDES CODANT LESDITS POLYPEPTIDES ET APPLICATIONS AU DIAGNOSTIC ET AU TRAITEMENT DU CANCER



(57) Abstract: The invention concerns a novel ICBP90 (Inverted CCAAT box binding protein 90) and its fragments, polynucleotides coding for said polypeptides and specific antibodies directed against said polypeptides. The invention also concerns methods and kits for diagnosing cell proliferation and compounds useful as medicine for preventing and/or treating pathology involving cell proliferation and in particular cancer.

(57) Abrégé: L'invention concerne un nouveau polypeptide ICBP90 (Inverted CCAAT box binding protein 90) et ses fragments, les polynucléotides codant pour lesdits polypeptides et des anticorps spécifiques dirigés contre lesdits polypeptides. L'invention concerne également des procédés et des kits de diagnostic de prolifération cellulaire et des composés utilisables à titre de médicament pour la prévention et/ou le traitement de pathologie faisant intervenir la prolifération cellulaire et du cancer en particulier.

WO 00/78949 A1

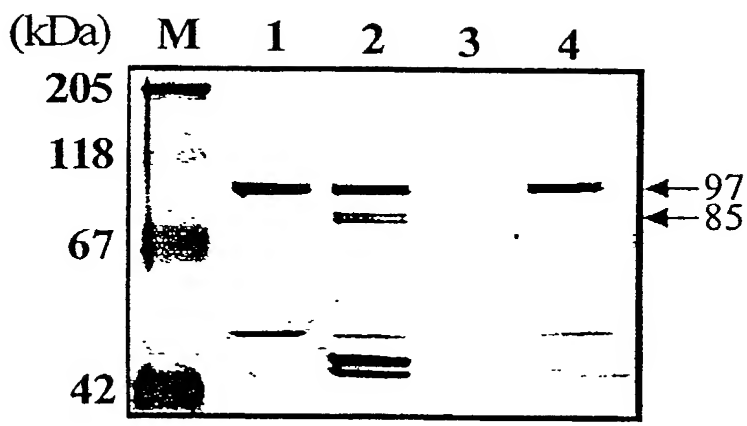


FIG_1

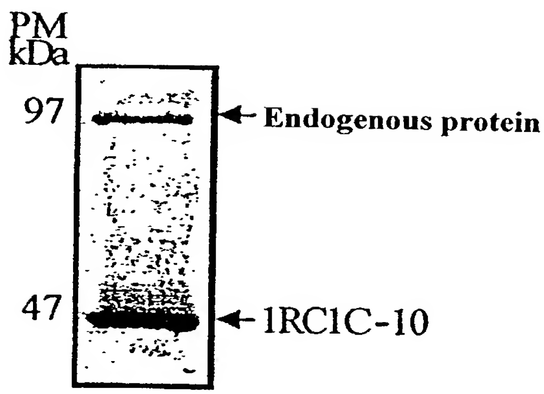


FIG_2

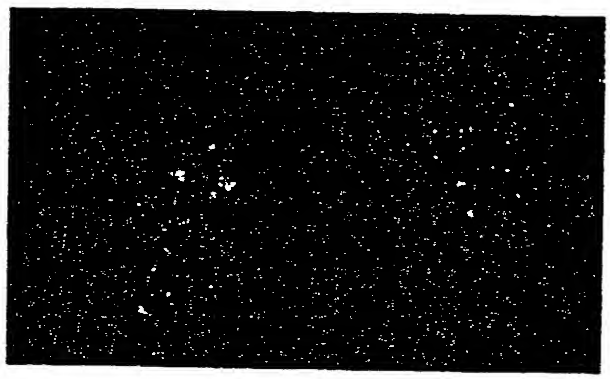
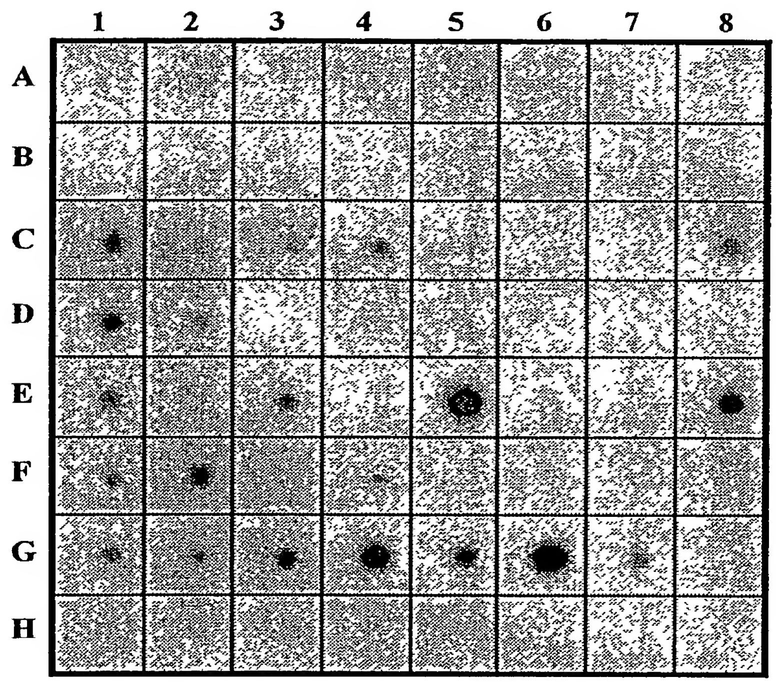


FIG_3

FIG_4



	1	2	3	4	5	6	7	8
A	whole brain	amygdala	caudate nucleus	cere-bellum	cerebral cortex	frontal lobe	hippo-campus	medulla oblongata
B	occipital lobe	putamen	substancia nigra	temporal lobe	thalamus	sub-thalamic nucleus	spinal cord	
C	heart	aorta	skeletal muscle	colon	bladder	uterus	prostate	stomach
D	testis	ovary	pancreas	pituitary gland	adrenal gland	thyroid gland	salivary gland	mammary gland
E	kidney	liver	small intestina	spleen	thymus	peripheral leukocyte	lymph node	bone marrow
F	appendix	lung	trachea	placenta				
G	fetal brain	fetal heart	fetal kidney	fetal liver	fetal spleen	fetal thymus	fetal lung	
H	yeast total RNA 100 ng	yeast tRNA 100 ng	E. coli rRNA 100 ng	E. coli DNA 100 ng	Poly r(A) 100 ng	human Cot1 DNA 100 ng	human DNA 100 ng	human DNA 500 ng

```
          |  10      |  20      |  30      |  40       |  50      |  60      |  70      |  80
   1 ATGTGGATCC AGGTTCGGAC CATGGATGGG AGGCAGACCC ACACGGTGGA CTCGCTGTCC AGGCTGACCA AGGTGGAGGA 80
  81 GCTGAGGCGG AAGATCCAGG AGCTGTTCCA CGTGGAGCCA GGCCTGCAGA GGCTGTTCTA CAGGGGCAAA CAGATGGAGG 160
 161 ACGGCCATAC CCTCTTCGAC TACGAGGTCC GCCTGAATGA CACCATCCAG CTCCTGGTCC GCCAGAGCCT CGTGCTCCCC 240
 241 CACAGCACCA AGGAGCGGGA CTCCGAGCTC TCCGACACCG ACTCCGGCTG CTGCCTGGGC CAGAGTGAGT CAGACAAGTC 320
 321 CTCCACCCAC GGTGAGGCGG CCGCCGAGAC TGACAGCAGG CCAGCCGATG AGGACATGTG GGATGAGACG GAATTGGGGC 400
 401 TGTACAAGGT CAATGAGTAC GTCGATGCTC GGGACACGAA CATGGGGGCG TGGTTTGAGG CGCAGGTGGT CAGGGTGACG 480
 481 CGGAAGGCCC CCTCCCGGGA CGAGCCCTGC AGCTCCACGT CCAGGCCGGC GCTGGAGGAG GACGTCATTT ACCACGTGAA 560
 561 ATACGACGAC TACCCGGAGA ACGGCGTGGT CCAGATGAAC TCCAGGGACG TCCGAGCGCG CGCCCGCACC ATCATCAAGT 640
 641 GGCAGGACCT GGAGGTGGGC CAGGTGGTCA TGCTCAACTA CAACCCCGAC AACCCCAAGG AGCGGGGCTT CTGGTACGAC 720
 721 GCGGAGATCT CCAGGAAGCG CGAGACCAGG ACGGCGCGGG AACTCTACGC CAACGTGGTG CTGGGGGATG ATTCTCTGAA 800
 801 CGACTGTCGG ATCATCTTCG TGGACGAAGT CTTCAAGATT GAGCGGCCGG GTGAAGGGAG CCCCATGGTT GACAACCCCA 880
 881 TGAGACGGAA GAGCGGGCCG TCCTGCAAGC ACTGCAAGGA CGACGTGAAC AGACTCTGCA GGGTCTGCGC CTGCCACCTG 960
 961 TGCGGGGGCC GGCAGGACCC CGACAAGCAG CTCATGTGCG ATGAGTGCGA CATGGCCTTC CACATCTACT GCCTGGACCC 1040
1041 GCCCCTCAGC AGTGTTCCCA GCGAGGACGA GTGGTACTGC CCTGAGTGCC GGAATGATGC CAGCGAGGTG GTACTGGCGG 1120
1121 GAGAGCGGCT GAGAGAGAGC AAGAAGAATG CGAAGATGGC CTCGGCCACA TCGTCCTCAC AGCGGGACTG GGGCAAGGGC 1200
1201 ATGGCCTGTG TGGGCCGCAC CAAGGAATGT ACCATCGTCC CGTCCAACCA CTACGGACCC ATCCCGGGGA TCCCCGTGGG 1280
1281 CACCATGTGG CGGTTCCGAG TCCAGGTCAG CGAGTCGGGT GTCCATCGGC CCCACGTGGC TGGCATCCAT GGCCGGAGCA 1360
1361 ACGACGGATC GTACTCCCTA GTCCTGGCGG GGGGCTATGA GGATGATGTG GACCATGGGA ATTTTTTCAC ATACACGGGT 1440
1441 AGTGGTGGTC GAGATCTTTC CGGCAACAAG AGGACCGCGG AACAGTCTTG TGATCAGAAA CTCACCAACA CCAACAGGGC 1520
1521 GCTGGCTCTC AACTGCTTTG CTCCCATCAA TGACCAAGAA GGGGCCGAGG CCAAGGACTG GCGGTCGGGG AAGCCGGTCA 1600
1601 GGGTGGTGCG CAATGTCAAG GGTGGCAAGA ATAGCAAGTA CGCCCCCGCT GAGGGCAACC GCTACGATGG CATCTACAAG 1680
1681 GTTGTGAAAT ACTGGCCCGA GAAGGGGAAG TCCGGGTTTC TCGTGTGGCG CTACCTTCTG CGGAGGGACG ATGATGAGCC 1760
1761 TGGCCCTTGG ACGAAGGAGG GGAAGGACCG GATCAAGAAG CTGGGGCTGA CCATGCAGTA TCCAGAAGGC TACCTGGAAG 1840
1841 CCCTGGCCAA CCGAGAGCGA GAGAAGGAGA ACAGCAAGAG GGAGGAGGAG GAGCAGCAGG AGGGGGGCTT CGCGTCCCCC 1920
1921 AGGACGGGCA AGGGCAAGTG GAAGCGGAAG TCGGCAGGAG GTGGCCCGAG CAGGGCCGGG TCCCCGCGCC GGACATCCAA 2000
2001 GAAAACCAAG GTGGAGCCCT ACAGTCTCAC GGCCCAGCAG AGCAGCCTCA TCAGAGAGGA CAAGAGCAAC GCCAAGCTGT 2080
2081 GGAATGAGGT CCTGGCGTCA CTCAAGGACC GGCCGGCGAG CGGCAGCCCG TTCCAGTTGT TCCTGAGTAA AGTGGAGGAG 2160
2161 ACGTTCCAGT GTATCTGCTG TCAGGAGCTG GTGTTCCGGC CCATCACGAC CGTGTGCCAG CACAACGTGT GCAAGGACTG 2240
2241 CCTGGACAGA TCCTTTCGGG CACAGGTGTT CAGCTGCCCT GCCTGCCGCT ACGACCTGGG CCGCAGCTAT GCCATGCAGG 2320
2321 TGAACCAGCC TCTGCAGACC GTCCTCAACC AGCTCTTCCC CGGCTACGGC AATGGCCGGT GA                    2382
          |  10      |  20      |  30      |  40       |  50      |  60      |  70      |  80
```

FIG_6

```
         |    10   |    20   |    30   |    40     |    50   |    60   |    70   |    80
  1 MWIQVRTMDG RQTHTVDSLS RLTKVEELRR KIQELFHVEP  GLQRLFYRGK QMEDGHTLFD YEVRLNDTIQ LLVRQSLVLP 80
 81 HSTKERDSEL SDTDSGCCLG QSESDKSSTH GEAAAETDSR  PADEDMWDET ELGLYKVNEY VDARDTNMGA WFEAQVVRVT 160
161 RKAPSRDEPC SSTSRPALEE DVIYHVKYDD YPENGVVQMN  SRDVRARART IIKWQDLEVG QVVMLNYNPD NPKERGFWYD 240
241 AEISRKRETR TARELYANVV LGDDSLNDCR IIFVDEVFKI  ERPGEGSPMV DNPMRRKSGP SCKHCKDDVN RLCRVCACHL 320
321 CGGRQDPDKQ LMCDECDMAF HIYCLDPPLS SVPSEDEWYC  PECRNDASEV VLAGERLRES KKNAKMASAT SSSQRDWGKG 400
401 MACVGRTKEC TIVPSNHYGP IPGIPVGTMW RFRVQVSESG  VHRPHVAGIH GRSNDGSYSL VLAGGYEDDV DHGNFFTYTG 480
481 SGGRDLSGNK RTAEQSCDQK LTNTNRALAL NCFAPINDQE  GAEAKDWRSG KPVRVVRNVK GGKNSKYAPA EGNRYDGIYK 560
561 VVKYWPEKGK SGFLVWRYLL RRDDDEPGPW TKEGKDRIKK  LGLTMQYPEG YLEALANRER EKENSKREEE EQQEGGFASP 640
641 RTGKGKWKRK SAGGGPSRAG SPRRTSKKTK VEPYSLTAQQ  SSLIREDKSN AKLWNEVLAS LKDRPASGSP FQLFLSKVEE 720
721 TFQCICCQEL VFRPITTVCQ HNVCKDCLDR SFRAQVFSCP  ACRYDLGRSY AMQVNQPLQT VLNQLFPGYG NGR*       794
         |    10   |    20   |    30   |    40     |    50   |    60   |    70   |    80
```

FIG_7

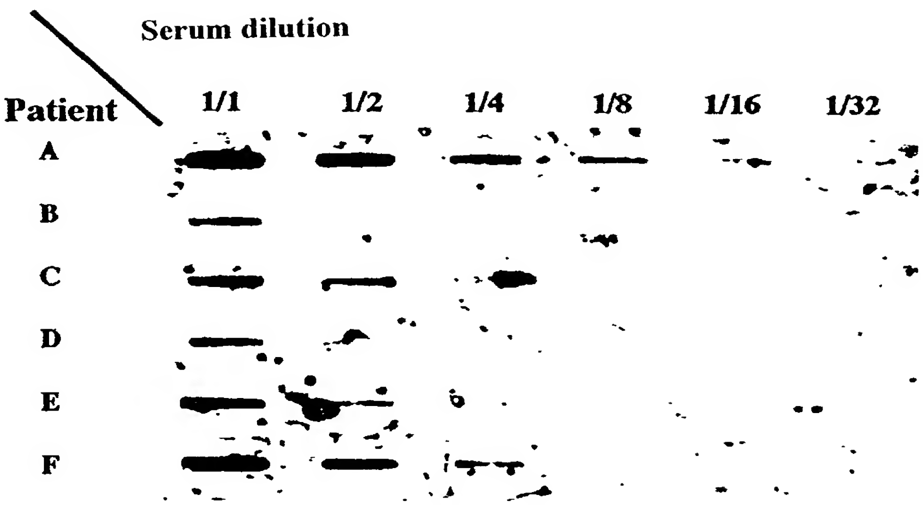
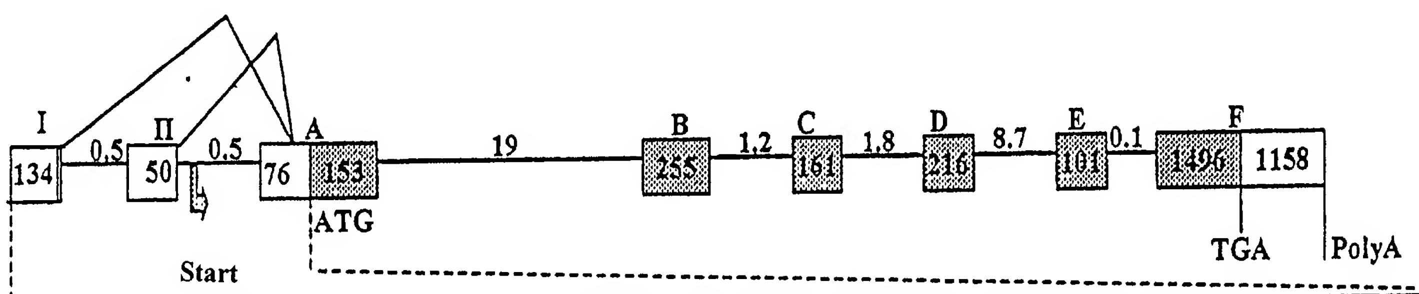


FIG-8

FIG-9

A



B

```
                                                          GC4
   1 GGCAGCGTTT GCCGAGCGGG CGCTCCGGGT CGCACGCAAG TCCGCGCGGG  GTCCGGGCCA CGCACGCGGT TTCATCGCCA TCCCCAGCCG GGCCAGGCGC
 101 GCAGGCAGAC AAGCTGTTCG CGGCGACCGG AGAGgtgagc gggcgggccg  ggtcggggtg ccagcccggg ccgggcgcac ggggctcggg aactttgcaa
 201 aactttcccg cgcggccagc ccgggcgcac gcatgtcccg cactctgtcc  cgggatccag ggcctcccct tccacctaac cctcgggaat cgttccccgg
 301 cacacatccg gctggagccg ggaccagcgc tgcgtccccg gagcccggcg  gggggtcgag cgcgccgggt gggggagggc ctggcgagcc gccggggagg
                                                                                                             CB2
 401 atgtcaggct ccgcgcctgc gcgcggggcg ccccgcgatt caattgtcgc  gcccgagccc gatttcgcgc gccctgagtt ccccgggagc atctgggcca
                 GC3                          CB1                    GC2
 501 atggggagcg agcggggcgg ggcggccggg tgctgcggag ccaataagag  gcggctcaag tgaagggggg cgggacttga cgagcggggg ccccctctgt
 601 agtcccggcg gcgggggtgg gcgtgggctc gctggcgcga cccgcgcggg  ccagtgggag tgcgggaggg acgccGAGGG TCCAGGGTTT GGAGGGGCGC
                                                                                                         GC1
 701 GAGCTGCCGG GGGTTGGAGG TCGAGgtgag tcgcggggcg cgcgcgctcg  cgggtggccg ggacggggcg cggttaccat ggccaccgcg gggcgggccc
 801 ggtcgcgcac gcgcgcgggg ggggccggca aggaggggggg gcgtgggcac  cgaggggtcc cggggtccgc ggatctcggg tggggttttt cccatttcag
 901 tggcacttgg ttaagttccc ccgggacctt ctgaagttcc ggcccgcgct  ggactttctg ggattccctc ttccgtaaat aggaatccga ggaatgaatg
1001 aatcaatgaa tgaatgaata aacgaaccaa ctcgggccac ttggcccggg  cctcctttct cctctggtcg tggggaagga gggatgggtt ggaccttctg
1101 cttttctttc aattccctct tttcattctc cttcctcctc aatcttcaac  acttggctag tcgttaatgc cttaagtgct taatttgttg tgtctggtcc
1201 tggccagggt ctggctgtac aggaggactg gaagggcatc ctgggagttt  cctggtgtcc acaggccgga caaaagcaac cccgactccT TAGAGCATGG
1301 CATGGCTCAG AGGTGCTGGT AAAACTGATG GGGGTTTATG CTGTCCCTCC  CCTCAGCGCC GACACCATG        1369
```

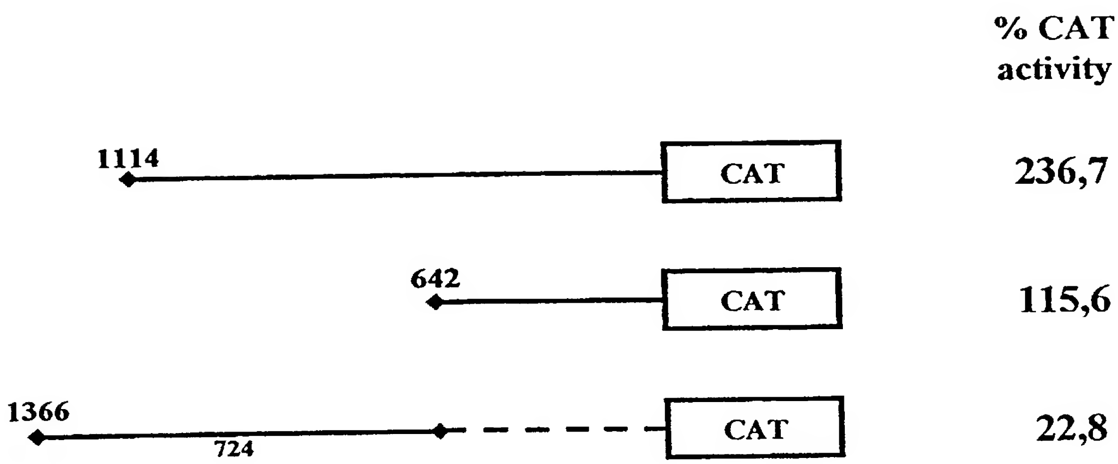


FIG 10

FIG-11A

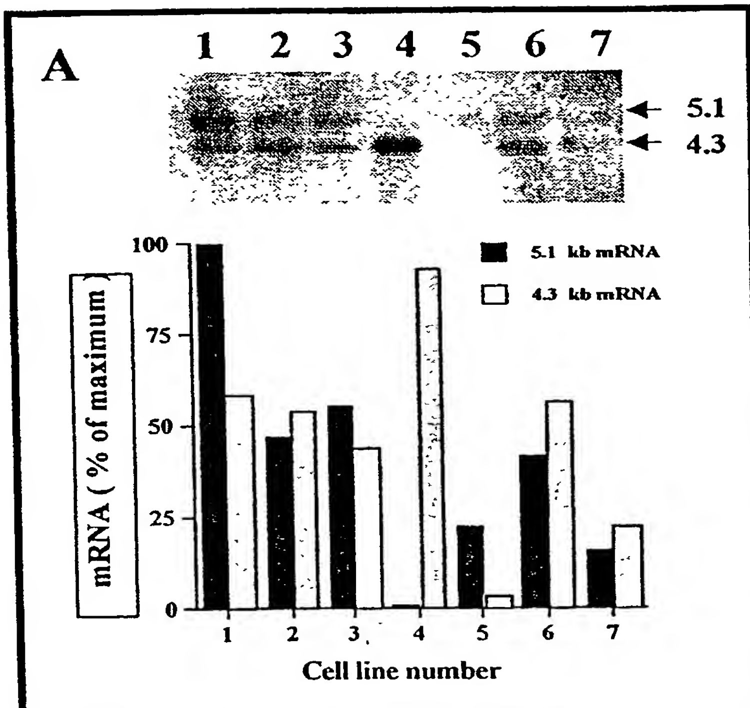
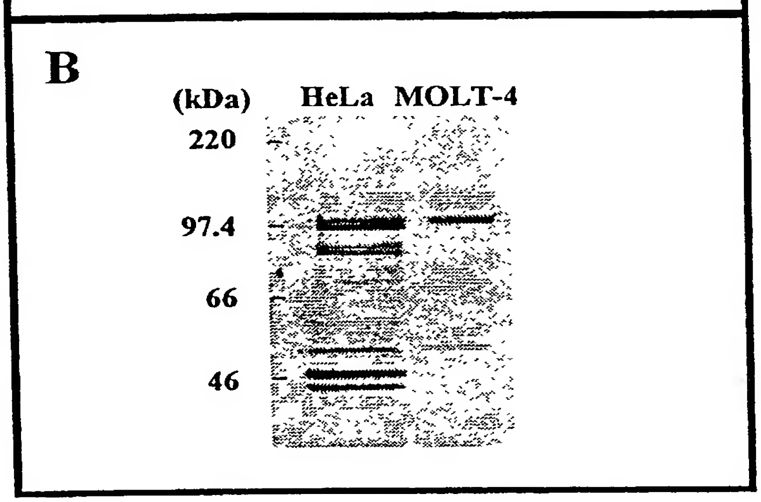


FIG-11B



10019071.053102

Docket No ____________________ #3

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:
My residence, post office address and citizenship are as stated below next to my name.
I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:
ICBP90 POLYPEPTIDE AND ITS FRAGMENTS AND POLYNUCLEOTIDES CODING FOR SAID POLYPEPTIDES AND APPLICATIONS TO THE DIAGNOSIS AND TREATMENT OF CANCER

The specification of which is attached hereto unless the following box is checked:
■ was filed on **June 26, 2000**____________ as (United States Application Number) or PCT International Application Number **PCT/FR00/01747**____________ and was amended on ____________ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.
I acknowledge the duty to disclose information which is known by me to be material to patentability as defined in Title 37, Code of Federal Regulations § 1.56.
I hereby claim foreign priority benefits under Title 35, United States Code, § 119(a)-(d) or § 365(b) of any foreign applications) for patent or inventor's certificate, or § 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below any foreign application for patent or inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed:

PRIOR FOREIGN APPLICATION(S)

NUMBER	COUNTRY	DAY/MONTH/YEAR FILED	PRIORITY CLAIMED
99 07935	FRANCE	22/06/99	yes

I hereby claim the benefit under Title 35, United States Code § 119(e) of any United States provisional applications) listed below.

APPLICATION NO.	FILING DATE

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s), or § 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose information which is known by me to be material to patentability as defined in Title 37, Code of Federal Regulations § 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

APPLICATION SERIAL NO.	FILING DATE	STATUS : PATENTED, PENDING ABANDONED

I hereby appoint as my attorneys, with full powers of substitution and revocation, to prosecute this application and transact all business in the Patent and Trademark Office connected therewith: Stephen A. Bent, Reg. No. 29,768; David A. Blumenthal Reg No 26 257-William T.Ellis, Reg. No. 26,874; John J. Feldhaus, Reg. No. 28,822; Patricia D. Granados, Reg. No. 33,683; John P. Isacson, Reg. No. 33.715; Donald D. Jeffery, Reg. No. 19,980; Eugene M. Lee. Reg. No. 32,039; Richard Linn, Reg. No.25,144; Peter G. Mack, Reg. No. 26,001; Brian J. McNamara, Reg. No. 32,789; Sybil Meloy. Reg. No. 22,749; George E. Quillin, Reg. No. 32,792; Colin G Sandercock, Reg. No. 31,298; Bernhard D. Saxe, Reg. No. 28,665; Charles F. Schill, Reg. No. 27.590; Richard L. Schwaab, Reg. No. 25,479; Arthur Schwartz, Reg. No. 22,115; Harold C. Wegner, Reg. No. 25,258.

19

Address all correspondence to FOLEY & LARDNER, Washington Harbour, 3000 K Street, N.W., Suite 500, P.O. Box 25696, Washington, D.C. 20007-8696. ~~Address telephone communications to~~ ~~at (202) 672-5300.~~

~~I hereby declare that~~ all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1-00

Full Name of First or Sole Inventor	Signature of First or Sole Inventor	Date
BRONNER Christian		February 11 2002
Residence Address	**Country of Citizenship**	
19, rue Exelmans 67640 FEGERSHEIM FRANCE FRX	FRANCE	
Post Office Address		
The same as residence		

2-00

Full Name of Second Inventor	Signature of Second Inventor	Date
HOPFNER Raphaël		February 11 2002
Residence Address	**Country of Citizenship**	
4, rue Fix 67000 STRASBOURG FRANCE FRX	FRANCE	
Post Office Address		
The same as residence		

3-00

Full Name of Third Inventor	Signature of Third Inventor	Date
MOUSLI Marc		February 11 2002
Residence Address	**Country of Citizenship**	
10, rue de Libreville 674000 ILLKIRCH FRANCE FRX	FRANCE	
Post Office Address		
The same as residence		

4-00

Full Name of Fourth Inventor	Signature of Fourth Inventor	Date
JELTSCH Jean-Marc		February 11 2002
Residence Address	**Country of Citizenship**	
1, rue de la Bruche 67120 MOLSHEIM FRANCE FRX	FRANCE	
Post Office Address		
The same as residence		

5-00

Full Name of Fifth Inventor	Signature of Fifth Inventor	Date
LUTZ Yves		February 11 2002
Residence Address	**Country of Citizenship**	
12, rue d'Ypres 67000 STRASBOURG FRANCE FRX	FRANCE	
Post Office Address		
The same as residence		

6.00

Full Name of Sixth Inventor	Signature of Sixth Inventor	Date
OUDET Pierre	P. Oudet	February 11 2002

Residence Address	Country of Citizenship
17, rue Vauban 67000 STRASBOURG FRANCE FRX	FRANCE

Post Office Address
The same as residence

Full Name of Seventh Inventor	Signature of Seventh Inventor	Date

Residence Address	Country of Citizenship

Post Office Address

Full Name of Eighth Inventor	Signature of Eighth Inventor	Date

Residence Address	Country of Citizenship

Post Office Address

Full Name of Ninth Inventor	Signature of Ninth Inventor	Date

Residence Address	Country of Citizenship

Post Office Address

Full Name of Tenth Inventor	Signature of Tenth Inventor	Date

Residence Address	Country of Citizenship

Post Office Address

Full Name of Eleventh Inventor	Signature of Eleventh Inventor	Date

Residence Address	Country of Citizenship

Post Office Address

SEQUENCE LISTING

<110> BRONNER Christian
HOPFNER Raphaël
MOUSLI Marc
JELTSCH Jean-Marc
LUTZ Yves
OUDET Pierre

<120> ICBP90 POLYPEPTIDE AND ITS FRAGMENTS AND POLYNUCLEOTIDES CODING FOR SAID POLYPEPTIDES AND APPLICATIONS FOR DIAGNOSING AND TREATING CANCER

<130> D18243

<150> FR 99 07935
<151> 1999-06-22

<150> PCT/FR00/01747
<151> 2000-06-22

<160> 12

<170> PatentIn Ver. 2.1

<210> 1
<211> 2382
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(2382)

<400> 1

```
atg tgg atc cag gtt cgg acc atg gat ggg agg cag acc cac acg gtg      48
Met Trp Ile Gln Val Arg Thr Met Asp Gly Arg Gln Thr His Thr Val
  1               5                  10                  15

gac tcg ctg tcc agg ctg acc aag gtg gag gag ctg agg cgg aag atc      96
Asp Ser Leu Ser Arg Leu Thr Lys Val Glu Glu Leu Arg Arg Lys Ile
             20                  25                  30

cag gag ctg ttc cac gtg gag cca ggc ctg cag agg ctg ttc tac agg     144
Gln Glu Leu Phe His Val Glu Pro Gly Leu Gln Arg Leu Phe Tyr Arg
         35                  40                  45

ggc aaa cag atg gag gac ggc cat acc ctc ttc gac tac gag gtc cgc     192
Gly Lys Gln Met Glu Asp Gly His Thr Leu Phe Asp Tyr Glu Val Arg
     50                  55                  60

ctg aat gac acc atc cag ctc ctg gtc cgc cag agc ctc gtg ctc ccc     240
Leu Asn Asp Thr Ile Gln Leu Leu Val Arg Gln Ser Leu Val Leu Pro
 65                  70                  75                  80

cac agc acc aag gag cgg gac tcc gag ctc tcc gac acc gac tcc ggc     288
His Ser Thr Lys Glu Arg Asp Ser Glu Leu Ser Asp Thr Asp Ser Gly
                 85                  90                  95

tgc tgc ctg ggc cag agt gag tca gac aag tcc tcc acc cac ggt gag     336
Cys Cys Leu Gly Gln Ser Glu Ser Asp Lys Ser Ser Thr His Gly Glu
```

```
          100                 105                 110

gcg gcc gcc gag act gac agc agg cca gcc gat gag gac atg tgg gat     384
Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp Glu Asp Met Trp Asp
        115                 120                 125

gag acg gaa ttg ggg ctg tac aag gtc aat gag tac gtc gat gct cgg     432
Glu Thr Glu Leu Gly Leu Tyr Lys Val Asn Glu Tyr Val Asp Ala Arg
    130                 135                 140

gac acg aac atg ggg gcg tgg ttt gag gcg cag gtg gtc agg gtg acg     480
Asp Thr Asn Met Gly Ala Trp Phe Glu Ala Gln Val Val Arg Val Thr
145                 150                 155                 160

cgg aag gcc ccc tcc cgg gac gag ccc tgc agc tcc acg tcc agg ccg     528
Arg Lys Ala Pro Ser Arg Asp Glu Pro Cys Ser Ser Thr Ser Arg Pro
                165                 170                 175

gcg ctg gag gag gac gtc att tac cac gtg aaa tac gac gac tac ccg     576
Ala Leu Glu Glu Asp Val Ile Tyr His Val Lys Tyr Asp Asp Tyr Pro
            180                 185                 190

gag aac ggc gtg gtc cag atg aac tcc agg gac gtc cga gcg cgc gcc     624
Glu Asn Gly Val Val Gln Met Asn Ser Arg Asp Val Arg Ala Arg Ala
        195                 200                 205

cgc acc atc atc aag tgg cag gac ctg gag gtg ggc cag gtg gtc atg     672
Arg Thr Ile Ile Lys Trp Gln Asp Leu Glu Val Gly Gln Val Val Met
    210                 215                 220

ctc aac tac aac ccc gac aac ccc aag gag cgg ggc ttc tgg tac gac     720
Leu Asn Tyr Asn Pro Asp Asn Pro Lys Glu Arg Gly Phe Trp Tyr Asp
225                 230                 235                 240

gcg gag atc tcc agg aag cgc gag acc agg acg gcg cgg gaa ctc tac     768
Ala Glu Ile Ser Arg Lys Arg Glu Thr Arg Thr Ala Arg Glu Leu Tyr
                245                 250                 255

gcc aac gtg gtg ctg ggg gat gat tct ctg aac gac tgt cgg atc atc     816
Ala Asn Val Val Leu Gly Asp Asp Ser Leu Asn Asp Cys Arg Ile Ile
            260                 265                 270

ttc gtg gac gaa gtc ttc aag att gag cgg ccg ggt gaa ggg agc ccc     864
Phe Val Asp Glu Val Phe Lys Ile Glu Arg Pro Gly Glu Gly Ser Pro
        275                 280                 285

atg gtt gac aac ccc atg aga cgg aag agc ggg ccg tcc tgc aag cac     912
Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
    290                 295                 300

tgc aag gac gac gtg aac aga ctc tgc agg gtc tgc gcc tgc cac ctg     960
Cys Lys Asp Asp Val Asn Arg Leu Cys Arg Val Cys Ala Cys His Leu
305                 310                 315                 320

tgc ggg ggc cgg cag gac ccc gac aag cag ctc atg tgc gat gag tgc    1008
Cys Gly Gly Arg Gln Asp Pro Asp Lys Gln Leu Met Cys Asp Glu Cys
                325                 330                 335

gac atg gcc ttc cac atc tac tgc ctg gac ccg ccc ctc agc agt gtt    1056
Asp Met Ala Phe His Ile Tyr Cys Leu Asp Pro Pro Leu Ser Ser Val
            340                 345                 350
```

```
ccc agc gag gac gag tgg tac tgc cct gag tgc cgg aat gat gcc agc     1104
Pro Ser Glu Asp Glu Trp Tyr Cys Pro Glu Cys Arg Asn Asp Ala Ser
        355                 360                 365

gag gtg gta ctg gcg gga gag cgg ctg aga gag agc aag aag aat gcg     1152
Glu Val Val Leu Ala Gly Glu Arg Leu Arg Glu Ser Lys Lys Asn Ala
    370                 375                 380

aag atg gcc tcg gcc aca tcg tcc tca cag cgg gac tgg ggc aag ggc     1200
Lys Met Ala Ser Ala Thr Ser Ser Ser Gln Arg Asp Trp Gly Lys Gly
385                 390                 395                 400

atg gcc tgt gtg ggc cgc acc aag gaa tgt acc atc gtc ccg tcc aac     1248
Met Ala Cys Val Gly Arg Thr Lys Glu Cys Thr Ile Val Pro Ser Asn
                405                 410                 415

cac tac gga ccc atc ccg ggg atc ccc gtg ggc acc atg tgg cgg ttc     1296
His Tyr Gly Pro Ile Pro Gly Ile Pro Val Gly Thr Met Trp Arg Phe
            420                 425                 430

cga gtc cag gtc agc gag tcg ggt gtc cat cgg ccc cac gtg gct ggc     1344
Arg Val Gln Val Ser Glu Ser Gly Val His Arg Pro His Val Ala Gly
        435                 440                 445

atc cat ggc cgg agc aac gac gga tcg tac tcc cta gtc ctg gcg ggg     1392
Ile His Gly Arg Ser Asn Asp Gly Ser Tyr Ser Leu Val Leu Ala Gly
    450                 455                 460

ggc tat gag gat gat gtg gac cat ggg aat ttt ttc aca tac acg ggt     1440
Gly Tyr Glu Asp Asp Val Asp His Gly Asn Phe Phe Thr Tyr Thr Gly
465                 470                 475                 480

agt ggt ggt cga gat ctt tcc ggc aac aag agg acc gcg gaa cag tct     1488
Ser Gly Gly Arg Asp Leu Ser Gly Asn Lys Arg Thr Ala Glu Gln Ser
                485                 490                 495

tgt gat cag aaa ctc acc aac acc aac agg gcg ctg gct ctc aac tgc     1536
Cys Asp Gln Lys Leu Thr Asn Thr Asn Arg Ala Leu Ala Leu Asn Cys
            500                 505                 510

ttt gct ccc atc aat gac caa gaa ggg gcc gag gcc aag gac tgg cgg     1584
Phe Ala Pro Ile Asn Asp Gln Glu Gly Ala Glu Ala Lys Asp Trp Arg
        515                 520                 525

tcg ggg aag ccg gtc agg gtg gtg cgc aat gtc aag ggt ggc aag aat     1632
Ser Gly Lys Pro Val Arg Val Val Arg Asn Val Lys Gly Gly Lys Asn
    530                 535                 540

agc aag tac gcc ccc gct gag ggc aac cgc tac gat ggc atc tac aag     1680
Ser Lys Tyr Ala Pro Ala Glu Gly Asn Arg Tyr Asp Gly Ile Tyr Lys
545                 550                 555                 560

gtt gtg aaa tac tgg ccc gag aag ggg aag tcc ggg ttt ctc gtg tgg     1728
Val Val Lys Tyr Trp Pro Glu Lys Gly Lys Ser Gly Phe Leu Val Trp
                565                 570                 575

cgc tac ctt ctg cgg agg gac gat gat gag cct ggc cct tgg acg aag     1776
Arg Tyr Leu Leu Arg Arg Asp Asp Asp Glu Pro Gly Pro Trp Thr Lys
            580                 585                 590
```

gag ggg aag gac cgg atc aag aag ctg ggg ctg acc atg cag tat cca 1824
Glu Gly Lys Asp Arg Ile Lys Lys Leu Gly Leu Thr Met Gln Tyr Pro
595 600 605

gaa ggc tac ctg gaa gcc ctg gcc aac cga gag cga gag aag gag aac 1872
Glu Gly Tyr Leu Glu Ala Leu Ala Asn Arg Glu Arg Glu Lys Glu Asn
610 615 620

agc aag agg gag gag gag gag cag cag gag ggg ggc ttc gcg tcc ccc 1920
Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly Gly Phe Ala Ser Pro
625 630 635 640

agg acg ggc aag ggc aag tgg aag cgg aag tcg gca gga ggt ggc ccg 1968
Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser Ala Gly Gly Gly Pro
645 650 655

agc agg gcc ggg tcc ccg cgc cgg aca tcc aag aaa acc aag gtg gag 2016
Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys Lys Thr Lys Val Glu
660 665 670

ccc tac agt ctc acg gcc cag cag agc agc ctc atc aga gag gac aag 2064
Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu Ile Arg Glu Asp Lys
675 680 685

agc aac gcc aag ctg tgg aat gag gtc ctg gcg tca ctc aag gac cgg 2112
Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala Ser Leu Lys Asp Arg
690 695 700

ccg gcg agc ggc agc ccg ttc cag ttg ttc ctg agt aaa gtg gag gag 2160
Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu Ser Lys Val Glu Glu
705 710 715 720

acg ttc cag tgt atc tgc tgt cag gag ctg gtg ttc cgg ccc atc acg 2208
Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val Phe Arg Pro Ile Thr
725 730 735

acc gtg tgc cag cac aac gtg tgc aag gac tgc ctg gac aga tcc ttt 2256
Thr Val Cys Gln His Asn Val Cys Lys Asp Cys Leu Asp Arg Ser Phe
740 745 750

cgg gca cag gtg ttc agc tgc cct gcc tgc cgc tac gac ctg ggc cgc 2304
Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg Tyr Asp Leu Gly Arg
755 760 765

agc tat gcc atg cag gtg aac cag cct ctg cag acc gtc ctc aac cag 2352
Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln Thr Val Leu Asn Gln
770 775 780

ctc ttc ccc ggc tac ggc aat ggc cgg tga 2382
Leu Phe Pro Gly Tyr Gly Asn Gly Arg
785 790

<210> 2
<211> 793
<212> PRT
<213> Homo sapiens

<400> 2
Met Trp Ile Gln Val Arg Thr Met Asp Gly Arg Gln Thr His Thr Val
1 5 10 15

```
Asp Ser Leu Ser Arg Leu Thr Lys Val Glu Glu Leu Arg Arg Lys Ile
            20                  25                  30

Gln Glu Leu Phe His Val Glu Pro Gly Leu Gln Arg Leu Phe Tyr Arg
        35                  40                  45

Gly Lys Gln Met Glu Asp Gly His Thr Leu Phe Asp Tyr Glu Val Arg
    50                  55                  60

Leu Asn Asp Thr Ile Gln Leu Leu Val Arg Gln Ser Leu Val Leu Pro
65                  70                  75                  80

His Ser Thr Lys Glu Arg Asp Ser Glu Leu Ser Asp Thr Asp Ser Gly
                85                  90                  95

Cys Cys Leu Gly Gln Ser Glu Ser Asp Lys Ser Ser Thr His Gly Glu
            100                 105                 110

Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp Glu Asp Met Trp Asp
        115                 120                 125

Glu Thr Glu Leu Gly Leu Tyr Lys Val Asn Glu Tyr Val Asp Ala Arg
    130                 135                 140

Asp Thr Asn Met Gly Ala Trp Phe Glu Ala Gln Val Val Arg Val Thr
145                 150                 155                 160

Arg Lys Ala Pro Ser Arg Asp Glu Pro Cys Ser Ser Thr Ser Arg Pro
                165                 170                 175

Ala Leu Glu Glu Asp Val Ile Tyr His Val Lys Tyr Asp Asp Tyr Pro
            180                 185                 190

Glu Asn Gly Val Val Gln Met Asn Ser Arg Asp Val Arg Ala Arg Ala
        195                 200                 205

Arg Thr Ile Ile Lys Trp Gln Asp Leu Glu Val Gly Gln Val Val Met
    210                 215                 220

Leu Asn Tyr Asn Pro Asp Asn Pro Lys Glu Arg Gly Phe Trp Tyr Asp
225                 230                 235                 240

Ala Glu Ile Ser Arg Lys Arg Glu Thr Arg Thr Ala Arg Glu Leu Tyr
                245                 250                 255

Ala Asn Val Val Leu Gly Asp Asp Ser Leu Asn Asp Cys Arg Ile Ile
            260                 265                 270

Phe Val Asp Glu Val Phe Lys Ile Glu Arg Pro Gly Glu Gly Ser Pro
        275                 280                 285

Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
    290                 295                 300

Cys Lys Asp Asp Val Asn Arg Leu Cys Arg Val Cys Ala Cys His Leu
305                 310                 315                 320

Cys Gly Gly Arg Gln Asp Pro Asp Lys Gln Leu Met Cys Asp Glu Cys
                325                 330                 335
```

Asp Met Ala Phe His Ile Tyr Cys Leu Asp Pro Pro Leu Ser Ser Val
340 345 350

Pro Ser Glu Asp Glu Trp Tyr Cys Pro Glu Cys Arg Asn Asp Ala Ser
355 360 365

Glu Val Val Leu Ala Gly Glu Arg Leu Arg Glu Ser Lys Lys Asn Ala
370 375 380

Lys Met Ala Ser Ala Thr Ser Ser Ser Gln Arg Asp Trp Gly Lys Gly
385 390 395 400

Met Ala Cys Val Gly Arg Thr Lys Glu Cys Thr Ile Val Pro Ser Asn
405 410 415

His Tyr Gly Pro Ile Pro Gly Ile Pro Val Gly Thr Met Trp Arg Phe
420 425 430

Arg Val Gln Val Ser Glu Ser Gly Val His Arg Pro His Val Ala Gly
435 440 445

Ile His Gly Arg Ser Asn Asp Gly Ser Tyr Ser Leu Val Leu Ala Gly
450 455 460

Gly Tyr Glu Asp Asp Val Asp His Gly Asn Phe Phe Thr Tyr Thr Gly
465 470 475 480

Ser Gly Gly Arg Asp Leu Ser Gly Asn Lys Arg Thr Ala Glu Gln Ser
485 490 495

Cys Asp Gln Lys Leu Thr Asn Thr Asn Arg Ala Leu Ala Leu Asn Cys
500 505 510

Phe Ala Pro Ile Asn Asp Gln Glu Gly Ala Glu Ala Lys Asp Trp Arg
515 520 525

Ser Gly Lys Pro Val Arg Val Val Arg Asn Val Lys Gly Gly Lys Asn
530 535 540

Ser Lys Tyr Ala Pro Ala Glu Gly Asn Arg Tyr Asp Gly Ile Tyr Lys
545 550 555 560

Val Val Lys Tyr Trp Pro Glu Lys Gly Lys Ser Gly Phe Leu Val Trp
565 570 575

Arg Tyr Leu Leu Arg Arg Asp Asp Asp Glu Pro Gly Pro Trp Thr Lys
580 585 590

Glu Gly Lys Asp Arg Ile Lys Lys Leu Gly Leu Thr Met Gln Tyr Pro
595 600 605

Glu Gly Tyr Leu Glu Ala Leu Ala Asn Arg Glu Arg Glu Lys Glu Asn
610 615 620

Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly Gly Phe Ala Ser Pro
625 630 635 640

Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser Ala Gly Gly Gly Pro
645 650 655

Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys Lys Thr Lys Val Glu

660 665 670

Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu Ile Arg Glu Asp Lys
675 680 685

Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala Ser Leu Lys Asp Arg
690 695 700

Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu Ser Lys Val Glu Glu
705 710 715 720

Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val Phe Arg Pro Ile Thr
725 730 735

Thr Val Cys Gln His Asn Val Cys Lys Asp Cys Leu Asp Arg Ser Phe
740 745 750

Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg Tyr Asp Leu Gly Arg
755 760 765

Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln Thr Val Leu Asn Gln
770 775 780

Leu Phe Pro Gly Tyr Gly Asn Gly Arg
785 790

<210> 3
<211> 45
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(45)

<400> 3
acc cac ggt gag gcg gcc gcc gag act gac agc agg cca gcc gat 45
Thr His Gly Glu Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp
1 5 10 15

<210> 4
<211> 15
<212> PRT
<213> Homo sapiens

<400> 4
Thr His Gly Glu Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp
1 5 10 15

<210> 5
<211> 78
<212> DNA
<213> Homo sapiens

<220>
<221> CDS

```
<222> (1)..(78)

<400> 5
atg gtt gac aac ccc atg aga cgg aag agc ggg ccg tcc tgc aag cac     48
Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
  1               5                  10                  15

tgc aag gac gac gtg aac aga ctc tgc agc                             78
Cys Lys Asp Asp Val Asn Arg Leu Cys Ser
             20                  25


<210> 6
<211> 26
<212> PRT
<213> Homo sapiens

<400> 6
Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
  1               5                  10                  15

Cys Lys Asp Asp Val Asn Arg Leu Cys Ser
             20                  25



<210> 7
<211> 525
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(525)

<400> 7
cga gag aag gag aac agc aag agg gag gag gag gag cag cag gag ggg     48
Arg Glu Lys Glu Asn Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly
  1               5                  10                  15

ggc ttc gcg tcc ccc agg acg ggc aag ggc aag tgg aag cgg aag tcg     96
Gly Phe Ala Ser Pro Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser
             20                  25                  30

gca gga ggt ggc ccg agc agg gcc ggg tcc ccg cgc cgg aca tcc aag    144
Ala Gly Gly Gly Pro Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys
         35                  40                  45

aaa acc aag gtg gag ccc tac agt ctc acg gcc cag cag agc agc ctc    192
Lys Thr Lys Val Glu Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu
     50                  55                  60

atc aga gag gac aag agc aac gcc aag ctg tgg aat gag gtc ctg gcg    240
Ile Arg Glu Asp Lys Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala
 65                  70                  75                  80

tca ctc aag gac cgg ccg gcg agc ggc agc ccg ttc cag ttg ttc ctg    288
Ser Leu Lys Asp Arg Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu
                 85                  90                  95

agt aaa gtg gag gag acg ttc cag tgt atc tgc tgt cag gag ctg gtg    336
```

```
Ser Lys Val Glu Glu Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val
            100                 105                 110

ttc cgg ccc atc acg acc gtg tgc cag cac aac gtg tgc aag gac tgc    384
Phe Arg Pro Ile Thr Thr Val Cys Gln His Asn Val Cys Lys Asp Cys
        115                 120                 125

ctg gac aga tcc ttt cgg gca cag gtg ttc agc tgc cct gcc tgc cgc    432
Leu Asp Arg Ser Phe Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg
    130                 135                 140

tac gac ctg ggc cgc agc tat gcc atg cag gtg aac cag cct ctg cag    480
Tyr Asp Leu Gly Arg Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln
145                 150                 155                 160

acc gtc ctc aac cag ctc ttc ccc ggc tac ggc aat ggc cgg tga        525
Thr Val Leu Asn Gln Leu Phe Pro Gly Tyr Gly Asn Gly Arg
                165                 170                 175


<210> 8
<211> 174
<212> PRT
<213> Homo sapiens

<400> 8
Arg Glu Lys Glu Asn Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly
  1               5                  10                  15

Gly Phe Ala Ser Pro Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser
             20                  25                  30

Ala Gly Gly Gly Pro Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys
        35                  40                  45

Lys Thr Lys Val Glu Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu
    50                  55                  60

Ile Arg Glu Asp Lys Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala
 65                  70                  75                  80

Ser Leu Lys Asp Arg Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu
                85                  90                  95

Ser Lys Val Glu Glu Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val
            100                 105                 110

Phe Arg Pro Ile Thr Thr Val Cys Gln His Asn Val Cys Lys Asp Cys
        115                 120                 125

Leu Asp Arg Ser Phe Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg
    130                 135                 140

Tyr Asp Leu Gly Arg Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln
145                 150                 155                 160

Thr Val Leu Asn Gln Leu Phe Pro Gly Tyr Gly Asn Gly Arg
                165                 170
```

<210> 9
<211> 324
<212> DNA
<213> Homo sapiens

<400> 9
atgtggatcc aggttcggac catggatggg aggcagaccc acacggtgga ctcgctgtcc 60
aggctgacca aggtggagga gctgaggcgg aagatccagg agctgttcca cgtggagcca 120
ggcctgcaga ggctgttcta caggggcaaa cagatggagg acggccatac cctcttcgac 180
tacgaggtcc gcctgaatga caccatccag ctcctggtcc gccagagcct cgtgctcccc 240
cacagcacca aggagcggga ctccgagctc tccgacaccg actccggctg ctgcctgggc 300
cagagtgagt cagacaagtc ctcc 324

<210> 10
<211> 495
<212> DNA
<213> Homo sapiens

<400> 10
gaggacatgt gggatgagac ggaattgggg ctgtacaagg tcaatgagta cgtcgatgct 60
cgggacacga acatgggggc gtggtttgag gcgcaggtgg tcagggtgac gcggaaggcc 120
ccctcccggg acgagccctg cagctccacg tccaggccgg cgctggagga ggacgtcatt 180
taccacgtga aatacgacga ctacccggag aacggcgtgg tccagatgaa ctccagggac 240
gtccgagcgc gcgcccgcac catcatcaag tggcaggacc tggaggtggg ccaggtggtc 300
atgctcaact acaaccccga caaccccaag gagcggggct tctggtacga cgcggagatc 360
tccaggaagc gcgagaccag gacggcgcgg gaactctacg ccaacgtggt gctgggggat 420
gattctctga acgactgtcg gatcatcttc gtggacgaag tcttcaagat tgagcggccg 480
ggtgaaggga gcccc 495

<210> 11
<211> 915
<212> DNA
<213> Homo sapiens

<400> 11
gtctgcgcct gccacctgtg cgggggccgg caggaccccg acaagcagct catgtgcgat 60
gagtgcgaca tggccttcca catctactgc ctggacccgc ccctcagcag tgttcccagc 120
gaggacgagt ggtactgccc tgagtgccgg aatgatgcca gcgaggtggt actggcggga 180
gagcggctga gagagagcaa gaagaatgcg aagatggcct cggccacatc gtcctcacag 240
cgggactggg gcaagggcat ggcctgtgtg ggccgcacca aggaatgtac catcgtcccg 300
tccaaccact acggacccat cccggggatc cccgtgggca ccatgtggcg gttccgagtc 360
caggtcagcg agtcgggtgt ccatcggccc cacgtggctg gcatccatgg ccggagcaac 420
gacggatcgt actccctagt cctggcgggg ggctatgagg atgatgtgga ccatgggaat 480
tttttcacat acacgggtag tggtggtcga gatctttccg gcaacaagag gaccgcggaa 540
cagtcttgtg atcagaaact caccaacacc aacaggcgc tggctctcaa ctgctttgct 600
cccatcaatg accaagaagg ggccgaggcc aaggactggc ggtcggggaa gccggtcagg 660
gtggtgcgca atgtcaaggg tggcaagaat agcaagtacg cccccgctga gggcaaccgc 720
tacgatggca tctacaaggt tgtgaaatac tggcccgaga aggggaagtc cgggtttctc 780
gtgtggcgct accttctgcg gagggacgat gatgagcctg gcccttggac gaaggagggg 840
aaggaccgga tcaagaagct ggggctgacc atgcagtatc cagaaggcta cctggaagcc 900
ctggccaacc gagag 915

<210> 12
<211> 1366
<212> DNA
<213> Homo sapiens

<400> 12

```
ggcagcgttt gccgagcggg cgctccgggt cgcacgcaag tccgcgcggg gtccgggcca 60
cgcacgcggt ttcatcgcca tccccagccg ggccaggcgc gcaggcagac aagctgttcg 120
cggcgaccgg agaggtgagc gggcgggccg ggtcggggtg ccagcccggg ccgggcgcac 180
ggggctcggg aactttgcaa aactttcccg cgcggccagc ccgggcgcac gcatgtcccg 240
cactctgtcc cgggatccag ggcctcccct tccacctaac cctcgggaat cgttccccgg 300
cacacatccg gctggagccg ggaccagcgc tgcgtccccg gagcccggcg gggggtcgag 360
cgcgccgggt gggggagggc ctggcgagcc gccggggagg atgtcaggct ccgcgcctgc 420
gcgcggggcg ccccgcgatt caattgtcgc gcccgagccc gatttcgcgc gccctgagtt 480
ccccgggagc atctgggcca atggggagcg agcggggcgg ggcggccggg tgctgcggag 540
ccaataagag gcggctcaag tgaagggggg cgggacttga cgagcggggg ccccctctgt 600
agtcccggcg gcgggggtgg gcgtgggctc gctggcgcga cccgcgcggg ccagtgggag 660
tgcgggaggg acgccgaggg tccagggttt ggaggggcgc gagctgccgg gggttggagg 720
tcgaggtgag tcgcggggcg cgcgcgctcg cgggtggccg ggacggggcg cggttaccat 780
ggccaccgcg gggcgggccc ggtcgcgcac gcgcgcgggg ggggccggca aggagggggg 840
gcgtgggcac cgaggggtcc cggggtccgc ggatctcggg tggggttttt cccatttcag 900
tggcacttgg ttaagttccc ccgggacctt ctgaagttcc ggcccgcgct ggactttctg 960
ggattccctc ttccgtaaat aggaatccga ggaatgaatg aatcaatgaa tgaatgaata 1020
aacgaaccaa ctcgggccac ttggcccggg cctcctttct cctctggtcg tggggaagga 1080
gggatgggtt ggaccttctg cttttctttc aattccctct tttcattctc cttcctcctc 1140
aatcttcaac acttggctag tcgttaatgc cttaagtgct taatttgttg tgtctggtcc 1200
tggccagggt ctggctgtac aggaggactg gaagggcatc ctgggagttt cctggtgtcc 1260
acaggccgga caaaagcaac cccgactcct tagagcatgg catggctcag aggtgctggt 1320
aaaactgatg gggtttatg ctgtccctcc cctcagcgcc gacacc 1366
```

LISTE DE SEQUENCES

<110> ADEREGEM

<120> Polypeptide ICBP90 et ses fragments et polynucléotides
codant lesdits polypeptides et applications au
diagnostic et au traitement du cancer

<130> d18243

<140>
<141>

<160> 12

<170> PatentIn Ver. 2.1

<210> 1
<211> 2382
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(2382)

<400> 1
atg tgg atc cag gtt cgg acc atg gat ggg agg cag acc cac acg gtg 48
Met Trp Ile Gln Val Arg Thr Met Asp Gly Arg Gln Thr His Thr Val
1 5 10 15

gac tcg ctg tcc agg ctg acc aag gtg gag gag ctg agg cgg aag atc 96
Asp Ser Leu Ser Arg Leu Thr Lys Val Glu Glu Leu Arg Arg Lys Ile
20 25 30

cag gag ctg ttc cac gtg gag cca ggc ctg cag agg ctg ttc tac agg 144
Gln Glu Leu Phe His Val Glu Pro Gly Leu Gln Arg Leu Phe Tyr Arg
35 40 45

ggc aaa cag atg gag gac ggc cat acc ctc ttc gac tac gag gtc cgc 192
Gly Lys Gln Met Glu Asp Gly His Thr Leu Phe Asp Tyr Glu Val Arg
50 55 60

ctg aat gac acc atc cag ctc ctg gtc cgc cag agc ctc gtg ctc ccc 240
Leu Asn Asp Thr Ile Gln Leu Leu Val Arg Gln Ser Leu Val Leu Pro
65 70 75 80

cac agc acc aag gag cgg gac tcc gag ctc tcc gac acc gac tcc ggc 288
His Ser Thr Lys Glu Arg Asp Ser Glu Leu Ser Asp Thr Asp Ser Gly
85 90 95

tgc tgc ctg ggc cag agt gag tca gac aag tcc tcc acc cac ggt gag 336
Cys Cys Leu Gly Gln Ser Glu Ser Asp Lys Ser Ser Thr His Gly Glu
100 105 110

gcg gcc gcc gag act gac agc agg cca gcc gat gag gac atg tgg gat 384
Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp Glu Asp Met Trp Asp
115 120 125

gag acg gaa ttg ggg ctg tac aag gtc aat gag tac gtc gat gct cgg 432
Glu Thr Glu Leu Gly Leu Tyr Lys Val Asn Glu Tyr Val Asp Ala Arg

```
130                  135                 140

gac acg aac atg ggg gcg tgg ttt gag gcg cag gtg gtc agg gtg acg    480
Asp Thr Asn Met Gly Ala Trp Phe Glu Ala Gln Val Val Arg Val Thr
145                 150                 155                 160

cgg aag gcc ccc tcc cgg gac gag ccc tgc agc tcc acg tcc agg ccg    528
Arg Lys Ala Pro Ser Arg Asp Glu Pro Cys Ser Ser Thr Ser Arg Pro
                165                 170                 175

gcg ctg gag gag gac gtc att tac cac gtg aaa tac gac gac tac ccg    576
Ala Leu Glu Glu Asp Val Ile Tyr His Val Lys Tyr Asp Asp Tyr Pro
            180                 185                 190

gag aac ggc gtg gtc cag atg aac tcc agg gac gtc cga gcg cgc gcc    624
Glu Asn Gly Val Val Gln Met Asn Ser Arg Asp Val Arg Ala Arg Ala
        195                 200                 205

cgc acc atc atc aag tgg cag gac ctg gag gtg ggc cag gtg gtc atg    672
Arg Thr Ile Ile Lys Trp Gln Asp Leu Glu Val Gly Gln Val Val Met
    210                 215                 220

ctc aac tac aac ccc gac aac ccc aag gag cgg ggc ttc tgg tac gac    720
Leu Asn Tyr Asn Pro Asp Asn Pro Lys Glu Arg Gly Phe Trp Tyr Asp
225                 230                 235                 240

gcg gag atc tcc agg aag cgc gag acc agg acg gcg cgg gaa ctc tac    768
Ala Glu Ile Ser Arg Lys Arg Glu Thr Arg Thr Ala Arg Glu Leu Tyr
                245                 250                 255

gcc aac gtg gtg ctg ggg gat gat tct ctg aac gac tgt cgg atc atc    816
Ala Asn Val Val Leu Gly Asp Asp Ser Leu Asn Asp Cys Arg Ile Ile
            260                 265                 270

ttc gtg gac gaa gtc ttc aag att gag cgg ccg ggt gaa ggg agc ccc    864
Phe Val Asp Glu Val Phe Lys Ile Glu Arg Pro Gly Glu Gly Ser Pro
        275                 280                 285

atg gtt gac aac ccc atg aga cgg aag agc ggg ccg tcc tgc aag cac    912
Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
    290                 295                 300

tgc aag gac gac gtg aac aga ctc tgc agg gtc tgc gcc tgc cac ctg    960
Cys Lys Asp Asp Val Asn Arg Leu Cys Arg Val Cys Ala Cys His Leu
305                 310                 315                 320

tgc ggg ggc cgg cag gac ccc gac aag cag ctc atg tgc gat gag tgc   1008
Cys Gly Gly Arg Gln Asp Pro Asp Lys Gln Leu Met Cys Asp Glu Cys
                325                 330                 335

gac atg gcc ttc cac atc tac tgc ctg gac ccg ccc ctc agc agt gtt   1056
Asp Met Ala Phe His Ile Tyr Cys Leu Asp Pro Pro Leu Ser Ser Val
            340                 345                 350

ccc agc gag gac gag tgg tac tgc cct gag tgc cgg aat gat gcc agc   1104
Pro Ser Glu Asp Glu Trp Tyr Cys Pro Glu Cys Arg Asn Asp Ala Ser
        355                 360                 365

gag gtg gta ctg gcg gga gag cgg ctg aga gag agc aag aag aat gcg   1152
Glu Val Val Leu Ala Gly Glu Arg Leu Arg Glu Ser Lys Lys Asn Ala
    370                 375                 380
```

aag atg gcc tcg gcc aca tcg tcc tca cag cgg gac tgg ggc aag ggc 1200
Lys Met Ala Ser Ala Thr Ser Ser Ser Gln Arg Asp Trp Gly Lys Gly
385 390 395 400

atg gcc tgt gtg ggc cgc acc aag gaa tgt acc atc gtc ccg tcc aac 1248
Met Ala Cys Val Gly Arg Thr Lys Glu Cys Thr Ile Val Pro Ser Asn
405 410 415

cac tac gga ccc atc ccg ggg atc ccc gtg ggc acc atg tgg cgg ttc 1296
His Tyr Gly Pro Ile Pro Gly Ile Pro Val Gly Thr Met Trp Arg Phe
420 425 430

cga gtc cag gtc agc gag tcg ggt gtc cat cgg ccc cac gtg gct ggc 1344
Arg Val Gln Val Ser Glu Ser Gly Val His Arg Pro His Val Ala Gly
435 440 445

atc cat ggc cgg agc aac gac gga tcg tac tcc cta gtc ctg gcg ggg 1392
Ile His Gly Arg Ser Asn Asp Gly Ser Tyr Ser Leu Val Leu Ala Gly
450 455 460

ggc tat gag gat gat gtg gac cat ggg aat ttt ttc aca tac acg ggt 1440
Gly Tyr Glu Asp Asp Val Asp His Gly Asn Phe Phe Thr Tyr Thr Gly
465 470 475 480

agt ggt ggt cga gat ctt tcc ggc aac aag agg acc gcg gaa cag tct 1488
Ser Gly Gly Arg Asp Leu Ser Gly Asn Lys Arg Thr Ala Glu Gln Ser
485 490 495

tgt gat cag aaa ctc acc aac acc aac agg gcg ctg gct ctc aac tgc 1536
Cys Asp Gln Lys Leu Thr Asn Thr Asn Arg Ala Leu Ala Leu Asn Cys
500 505 510

ttt gct ccc atc aat gac caa gaa ggg gcc gag gcc aag gac tgg cgg 1584
Phe Ala Pro Ile Asn Asp Gln Glu Gly Ala Glu Ala Lys Asp Trp Arg
515 520 525

tcg ggg aag ccg gtc agg gtg gtg cgc aat gtc aag ggt ggc aag aat 1632
Ser Gly Lys Pro Val Arg Val Val Arg Asn Val Lys Gly Gly Lys Asn
530 535 540

agc aag tac gcc ccc gct gag ggc aac cgc tac gat ggc atc tac aag 1680
Ser Lys Tyr Ala Pro Ala Glu Gly Asn Arg Tyr Asp Gly Ile Tyr Lys
545 550 555 560

gtt gtg aaa tac tgg ccc gag aag ggg aag tcc ggg ttt ctc gtg tgg 1728
Val Val Lys Tyr Trp Pro Glu Lys Gly Lys Ser Gly Phe Leu Val Trp
565 570 575

cgc tac ctt ctg cgg agg gac gat gat gag cct ggc cct tgg acg aag 1776
Arg Tyr Leu Leu Arg Arg Asp Asp Asp Glu Pro Gly Pro Trp Thr Lys
580 585 590

gag ggg aag gac cgg atc aag aag ctg ggg ctg acc atg cag tat cca 1824
Glu Gly Lys Asp Arg Ile Lys Lys Leu Gly Leu Thr Met Gln Tyr Pro
595 600 605

gaa ggc tac ctg gaa gcc ctg gcc aac cga gag cga gag aag gag aac 1872
Glu Gly Tyr Leu Glu Ala Leu Ala Asn Arg Glu Arg Glu Lys Glu Asn
610 615 620

```
agc aag agg gag gag gag gag cag cag gag ggg ggc ttc gcg tcc ccc     1920
Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly Gly Phe Ala Ser Pro
625                 630                 635                 640

agg acg ggc aag ggc aag tgg aag cgg aag tcg gca gga ggt ggc ccg     1968
Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser Ala Gly Gly Gly Pro
                645                 650                 655

agc agg gcc ggg tcc ccg cgc cgg aca tcc aag aaa acc aag gtg gag     2016
Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys Lys Thr Lys Val Glu
            660                 665                 670

ccc tac agt ctc acg gcc cag cag agc agc ctc atc aga gag gac aag     2064
Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu Ile Arg Glu Asp Lys
        675                 680                 685

agc aac gcc aag ctg tgg aat gag gtc ctg gcg tca ctc aag gac cgg     2112
Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala Ser Leu Lys Asp Arg
    690                 695                 700

ccg gcg agc ggc agc ccg ttc cag ttg ttc ctg agt aaa gtg gag gag     2160
Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu Ser Lys Val Glu Glu
705                 710                 715                 720

acg ttc cag tgt atc tgc tgt cag gag ctg gtg ttc cgg ccc atc acg     2208
Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val Phe Arg Pro Ile Thr
                725                 730                 735

acc gtg tgc cag cac aac gtg tgc aag gac tgc ctg gac aga tcc ttt     2256
Thr Val Cys Gln His Asn Val Cys Lys Asp Cys Leu Asp Arg Ser Phe
            740                 745                 750

cgg gca cag gtg ttc agc tgc cct gcc tgc cgc tac gac ctg ggc cgc     2304
Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg Tyr Asp Leu Gly Arg
        755                 760                 765

agc tat gcc atg cag gtg aac cag cct ctg cag acc gtc ctc aac cag     2352
Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln Thr Val Leu Asn Gln
    770                 775                 780

ctc ttc ccc ggc tac ggc aat ggc cgg tga                             2382
Leu Phe Pro Gly Tyr Gly Asn Gly Arg
785                 790


&lt;210&gt; 2
&lt;211&gt; 793
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;400&gt; 2
Met Trp Ile Gln Val Arg Thr Met Asp Gly Arg Gln Thr His Thr Val
1               5                   10                  15

Asp Ser Leu Ser Arg Leu Thr Lys Val Glu Glu Leu Arg Arg Lys Ile
            20                  25                  30

Gln Glu Leu Phe His Val Glu Pro Gly Leu Gln Arg Leu Phe Tyr Arg
        35                  40                  45

Gly Lys Gln Met Glu Asp Gly His Thr Leu Phe Asp Tyr Glu Val Arg
```

50 55 60

Leu Asn Asp Thr Ile Gln Leu Leu Val Arg Gln Ser Leu Val Leu Pro
65 70 75 80

His Ser Thr Lys Glu Arg Asp Ser Glu Leu Ser Asp Thr Asp Ser Gly
85 90 95

Cys Cys Leu Gly Gln Ser Glu Ser Asp Lys Ser Ser Thr His Gly Glu
100 105 110

Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp Glu Asp Met Trp Asp
115 120 125

Glu Thr Glu Leu Gly Leu Tyr Lys Val Asn Glu Tyr Val Asp Ala Arg
130 135 140

Asp Thr Asn Met Gly Ala Trp Phe Glu Ala Gln Val Val Arg Val Thr
145 150 155 160

Arg Lys Ala Pro Ser Arg Asp Glu Pro Cys Ser Ser Thr Ser Arg Pro
165 170 175

Ala Leu Glu Glu Asp Val Ile Tyr His Val Lys Tyr Asp Asp Tyr Pro
180 185 190

Glu Asn Gly Val Val Gln Met Asn Ser Arg Asp Val Arg Ala Arg Ala
195 200 205

Arg Thr Ile Ile Lys Trp Gln Asp Leu Glu Val Gly Gln Val Val Met
210 215 220

Leu Asn Tyr Asn Pro Asp Asn Pro Lys Glu Arg Gly Phe Trp Tyr Asp
225 230 235 240

Ala Glu Ile Ser Arg Lys Arg Glu Thr Arg Thr Ala Arg Glu Leu Tyr
245 250 255

Ala Asn Val Val Leu Gly Asp Asp Ser Leu Asn Asp Cys Arg Ile Ile
260 265 270

Phe Val Asp Glu Val Phe Lys Ile Glu Arg Pro Gly Glu Gly Ser Pro
275 280 285

Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
290 295 300

Cys Lys Asp Asp Val Asn Arg Leu Cys Arg Val Cys Ala Cys His Leu
305 310 315 320

Cys Gly Gly Arg Gln Asp Pro Asp Lys Gln Leu Met Cys Asp Glu Cys
325 330 335

Asp Met Ala Phe His Ile Tyr Cys Leu Asp Pro Pro Leu Ser Ser Val
340 345 350

Pro Ser Glu Asp Glu Trp Tyr Cys Pro Glu Cys Arg Asn Asp Ala Ser
355 360 365

Glu Val Val Leu Ala Gly Glu Arg Leu Arg Glu Ser Lys Lys Asn Ala
370 375 380

```
Lys Met Ala Ser Ala Thr Ser Ser Ser Gln Arg Asp Trp Gly Lys Gly
385                 390                 395                 400

Met Ala Cys Val Gly Arg Thr Lys Glu Cys Thr Ile Val Pro Ser Asn
                405                 410                 415

His Tyr Gly Pro Ile Pro Gly Ile Pro Val Gly Thr Met Trp Arg Phe
            420                 425                 430

Arg Val Gln Val Ser Glu Ser Gly Val His Arg Pro His Val Ala Gly
        435                 440                 445

Ile His Gly Arg Ser Asn Asp Gly Ser Tyr Ser Leu Val Leu Ala Gly
    450                 455                 460

Gly Tyr Glu Asp Asp Val Asp His Gly Asn Phe Phe Thr Tyr Thr Gly
465                 470                 475                 480

Ser Gly Gly Arg Asp Leu Ser Gly Asn Lys Arg Thr Ala Glu Gln Ser
                485                 490                 495

Cys Asp Gln Lys Leu Thr Asn Thr Asn Arg Ala Leu Ala Leu Asn Cys
            500                 505                 510

Phe Ala Pro Ile Asn Asp Gln Glu Gly Ala Glu Ala Lys Asp Trp Arg
        515                 520                 525

Ser Gly Lys Pro Val Arg Val Val Arg Asn Val Lys Gly Gly Lys Asn
    530                 535                 540

Ser Lys Tyr Ala Pro Ala Glu Gly Asn Arg Tyr Asp Gly Ile Tyr Lys
545                 550                 555                 560

Val Val Lys Tyr Trp Pro Glu Lys Gly Lys Ser Gly Phe Leu Val Trp
                565                 570                 575

Arg Tyr Leu Leu Arg Arg Asp Asp Asp Glu Pro Gly Pro Trp Thr Lys
            580                 585                 590

Glu Gly Lys Asp Arg Ile Lys Lys Leu Gly Leu Thr Met Gln Tyr Pro
        595                 600                 605

Glu Gly Tyr Leu Glu Ala Leu Ala Asn Arg Glu Arg Glu Lys Glu Asn
    610                 615                 620

Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly Gly Phe Ala Ser Pro
625                 630                 635                 640

Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser Ala Gly Gly Gly Pro
                645                 650                 655

Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys Lys Thr Lys Val Glu
            660                 665                 670

Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu Ile Arg Glu Asp Lys
        675                 680                 685

Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala Ser Leu Lys Asp Arg
    690                 695                 700
```

Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu Ser Lys Val Glu Glu
705 710 715 720

Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val Phe Arg Pro Ile Thr
725 730 735

Thr Val Cys Gln His Asn Val Cys Lys Asp Cys Leu Asp Arg Ser Phe
740 745 750

Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg Tyr Asp Leu Gly Arg
755 760 765

Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln Thr Val Leu Asn Gln
770 775 780

Leu Phe Pro Gly Tyr Gly Asn Gly Arg
785 790

<210> 3
<211> 45
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(45)

<400> 3
acc cac ggt gag gcg gcc gcc gag act gac agc agg cca gcc gat 45
Thr His Gly Glu Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp
1 5 10 15

<210> 4
<211> 15
<212> PRT
<213> Homo sapiens

<400> 4
Thr His Gly Glu Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp
1 5 10 15

<210> 5
<211> 78
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(78)

<400> 5
atg gtt gac aac ccc atg aga cgg aag agc ggg ccg tcc tgc aag cac 48
Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
1 5 10 15

tgc aag gac gac gtg aac aga ctc tgc agc 78

Cys Lys Asp Asp Val Asn Arg Leu Cys Ser
20 25

<210> 6
<211> 26
<212> PRT
<213> Homo sapiens

<400> 6
Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
1 5 10 15

Cys Lys Asp Asp Val Asn Arg Leu Cys Ser
20 25

<210> 7
<211> 525
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(525)

<400> 7
cga gag aag gag aac agc aag agg gag gag gag gag cag cag gag ggg 48
Arg Glu Lys Glu Asn Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly
1 5 10 15

ggc ttc gcg tcc ccc agg acg ggc aag ggc aag tgg aag cgg aag tcg 96
Gly Phe Ala Ser Pro Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser
20 25 30

gca gga ggt ggc ccg agc agg gcc ggg tcc ccg cgc cgg aca tcc aag 144
Ala Gly Gly Gly Pro Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys
35 40 45

aaa acc aag gtg gag ccc tac agt ctc acg gcc cag cag agc agc ctc 192
Lys Thr Lys Val Glu Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu
50 55 60

atc aga gag gac aag agc aac gcc aag ctg tgg aat gag gtc ctg gcg 240
Ile Arg Glu Asp Lys Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala
65 70 75 80

tca ctc aag gac cgg ccg gcg agc ggc agc ccg ttc cag ttg ttc ctg 288
Ser Leu Lys Asp Arg Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu
85 90 95

agt aaa gtg gag gag acg ttc cag tgt atc tgc tgt cag gag ctg gtg 336
Ser Lys Val Glu Glu Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val
100 105 110

ttc cgg ccc atc acg acc gtg tgc cag cac aac gtg tgc aag gac tgc 384
Phe Arg Pro Ile Thr Thr Val Cys Gln His Asn Val Cys Lys Asp Cys
115 120 125

ctg gac aga tcc ttt cgg gca cag gtg ttc agc tgc cct gcc tgc cgc 432

```
Leu Asp Arg Ser Phe Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg
    130                 135                 140

tac gac ctg ggc cgc agc tat gcc atg cag gtg aac cag cct ctg cag      480
Tyr Asp Leu Gly Arg Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln
145                 150                 155                 160

acc gtc ctc aac cag ctc ttc ccc ggc tac ggc aat ggc cgg tga          525
Thr Val Leu Asn Gln Leu Phe Pro Gly Tyr Gly Asn Gly Arg
                165                 170                 175


<210> 8
<211> 174
<212> PRT
<213> Homo sapiens

<400> 8
Arg Glu Lys Glu Asn Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly
  1               5                  10                  15

Gly Phe Ala Ser Pro Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser
             20                  25                  30

Ala Gly Gly Gly Pro Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys
         35                  40                  45

Lys Thr Lys Val Glu Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu
     50                  55                  60

Ile Arg Glu Asp Lys Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala
 65                  70                  75                  80

Ser Leu Lys Asp Arg Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu
                 85                  90                  95

Ser Lys Val Glu Glu Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val
            100                 105                 110

Phe Arg Pro Ile Thr Thr Val Cys Gln His Asn Val Cys Lys Asp Cys
        115                 120                 125

Leu Asp Arg Ser Phe Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg
    130                 135                 140

Tyr Asp Leu Gly Arg Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln
145                 150                 155                 160

Thr Val Leu Asn Gln Leu Phe Pro Gly Tyr Gly Asn Gly Arg
                165                 170


<210> 9
<211> 324
<212> ADN
<213> Homo sapiens

<400> 9
atgtggatcc aggttcggac catggatggg aggcagaccc acacggtgga ctcgctgtcc 60
aggctgacca aggtggagga gctgaggcgg aagatccagg agctgttcca cgtggagcca 120
```

ggcctgcaga ggctgttcta caggggcaaa cagatggagg acggccatac cctcttcgac 180
tacgaggtcc gcctgaatga caccatccag ctcctggtcc gccagagcct cgtgctcccc 240
cacagcacca aggagcggga ctccgagctc tccgacaccg actccggctg ctgcctgggc 300
cagagtgagt cagacaagtc ctcc 324

<210> 10
<211> 495
<212> ADN
<213> Homo sapiens

<400> 10
gaggacatgt gggatgagac ggaattgggg ctgtacaagg tcaatgagta cgtcgatgct 60
cgggacacga acatgggggc gcggtttgag gcgcaggtgg tcagggtgac gcggaaggcc 120
ccctcccggg acgagccctg cagctccacg tccaggccgg cgctggagga ggacgtcatt 180
taccacgtga aatacgacga ctacccggag aacggcgtgg tccagatgaa ctccagggac 240
gtccgagcgc gcgcccgcac catcatcaag tggcaggacc tggaggtggg ccaggtggtc 300
atgctcaact acaaccccga caaccccaag gagcggggct tctggtacga cgcggagatc 360
tccaggaagc gcgagaccag gacggcgcgg gaactctacg ccaacgtggt gctgggggat 420
gattctctga acgactgtcg gatcatcttc gtggacgaag tcttcaagat tgagcggccg 480
ggtgaaggga gcccc 495

<210> 11
<211> 915
<212> ADN
<213> Homo sapiens

<400> 11
gtctgcgcct gccacctgtg cgggggccgg caggaccccg acaagcagct catgtgcgat 60
gagtgcgaca tggccttcca catctactgc ctggacccgc ccctcagcag tgttcccagc 120
gaggacgagt ggtactgccc tgagtgccgg aatgatgcca gcgaggtggt actggcggga 180
gagcggctga gagagagcaa gaagaatgcg aagatggcct cggccacatc gtcctcacag 240
cgggactggg gcaagggcat ggcctgtgtg ggccgcacca aggaatgtac catcgtcccg 300
tccaaccact acggacccat cccggggatc cccgtgggca ccatgtggcg gttccgagtc 360
caggtcagcg agtcgggtgt ccatcggccc cacgtggctg gcatccatgg ccggagcaac 420
gacggatcgt actccctagt cctggcgggg ggctatgagg atgatgtgga ccatgggaat 480
tttttcacat acacgggtag tggtggtcga gatctttccg gcaacaagag gaccgcggaa 540
cagtcttgtg atcagaaact caccaacacc aacagggcgc tggctctcaa ctgctttgct 600
cccatcaatg accaagaagg ggccgaggcc aaggactggc ggtcggggaa gccggtcagg 660
gtggtgcgca atgtcaaggg tggcaagaat agcaagtacg cccccgctga gggcaaccgc 720
tacgatggca tctacaaggt tgtgaaatac tggcccgaga aggggaagtc cgggtttctc 780
gtgtggcgct accttctgcg gagggacgat gatgagcctg gcccttggac gaaggagggg 840
aaggaccgga tcaagaagct ggggctgacc atgcagtatc cagaaggcta cctggaagcc 900
ctggccaacc gagag 915

<210> 12
<211> 1366
<212> ADN
<213> Homo sapiens

<400> 12
ggcagcgttt gccgagcggg cgctccgggt cgcacgcaag tccgcgcggg gtccgggcca 60
cgcacgcggt ttcatcgcca tccccagccg ggccaggcgc gcaggcagac aagctgttcg 120
cggcgaccgg agaggtgagc gggcgggccg ggtcggggtg ccagcccggg ccgggcgcac 180
ggggctcggg aactttgcaa aactttcccg cgcggccagc ccgggcgcac gcatgtcccg 240
cactctgtcc cgggatccag ggcctcccct tccacctaac cctcgggaat cgttccccgg 300
cacacatccg gctggagccg ggaccagcgc tgcgtccccg gagcccggcg gggggtcgag 360
cgcgccgggt gggggagggc ctggcgagcc gccggggagg atgtcaggct ccgcgcctgc 420
gcgcggggcg ccccgcgatt caattgtcgc gcccgagccc gatttcgcgc gccctgagtt 480

```
ccccgggagc atctgggcca atggggagcg agcggggcgg ggcggccggg tgctgcggag 540
ccaataagag gcggctcaag tgaagggggg cgggacttga cgagcggggg ccccctctgt 600
agtcccggcg gcgggggtgg gcgtgggctc gctggcgcga cccgcgcggg ccagtgggag 660
tgcgggaggg acgccgaggg tccagggttt ggagggggcgc gagctgccgg gggttggagg 720
tcgaggtgag tcgcggggcg cgcgcgctcg cgggtggccg ggacggggcg cggttaccat 780
ggccaccgcg gggcgggccc ggtcgcgcac gcgcgcgggg ggggccggca aggagggggg 840
gcgtgggcac cgaggggtcc cggggtccgc ggatctcggg tggggttttt cccatttcag 900
tggcacttgg ttaagttccc ccgggacctt ctgaagttcc ggcccgcgct ggactttctg 960
ggattccctc ttccgtaaat aggaatccga ggaatgaatg aatcaatgaa tgaatgaata 1020
aacgaaccaa ctcgggccac ttggcccggg cctcctttct cctctggtcg tggggaagga 1080
gggatgggtt ggaccttctg cttttctttc aattccctct tttcattctc cttcctcctc 1140
aatcttcaac acttggctag tcgttaatgc cttaagtgct taatttgttg tgtctggtcc 1200
tggccagggt ctggctgtac aggaggactg gaagggcatc ctgggagttt cctggtgtcc 1260
acaggccgga caaaagcaac cccgactcct tagagcatgg catggctcag aggtgctggt 1320
aaaactgatg gggtttatg ctgtccctcc cctcagcgcc gacacc 1366
```